Author:
Saunders, Samantha J

Title:
Establishing a Reverse Genetic System for Type 1 Feline Coronavirus

Establishing a Reverse Genetic System for Type 1 Feline Coronavirus

Samantha Jane Saunders

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Life Sciences, School of Cellular and Molecular Medicine, on 20th September 2019

Word count (excluding preliminary pages, references and appendices): 52,547

Abstract

Feline coronavirus (FCoV) leads to the fatal disease known as feline infectious peritonitis in a small proportion of infected cats. Research into FCoV has so far been hindered by our inability to culture the most common serotype, Type 1, *in vitro*.

This project aimed to establish a reverse genetic system for Type 1 FCoV. Such a system would enable a deeper understanding of the role of viral mutations in the pathogenesis of disease and could serve as a platform for rational vaccine design. To this end, cDNA infectious clone and replicon constructs based on Type 1 FCoV were developed and transcribed. During this project, infectious virus or replicon-expressing cells were not recovered following transfection of the constructs into mammalian cells, but a foundation was laid for this to be achieved in future.

Recovery of recombinant Type 1 FCoV would necessitate a cell line capable of supporting this serotype's growth *in vitro*, so an aim of this project was to identify a cell entry receptor (CER) for Type 1 FCoV in order to develop a cell line permissive to infection with the virus. To this end, 'bait proteins' bearing the spike proteins of Type 1 and 2 FCoV were produced, and Type 2 bait protein was able to recognise its CER. Feline intestinal organoid cultures were established, and an interaction identified between Type 1 bait protein and heat shock 70 kDa protein 1a (HSPA1A) suggested that HSPA1A is a receptor for Type 1 FCoV.

Finally, a feline IFN-γ ELISpot assay was used to measure the cellular immune response to peptides representing a Type 1 FCoV epitope with and without a substitution. No significant difference in the ability of the peptide variants to stimulate a response was identified, but the assay could be used in the future to guide rational vaccine design.

Acknowledgments

I would like to thank my supervisors, Andrew Davidson and Séverine Tasker, for their wisdom, support and patience throughout this process. Special thanks go to the members of the E50 virology lab for their help, encouragement and friendship. I would also like to acknowledge Linda Wooldridge and Anya Lissina for supervising the ELISpot part of this project and, with others in Linda's group, giving me an excellent introduction into the weird world of postgraduate research.

Thanks to all the wonderful scientists who contributed their time and expertise to this project, including Anja Kipar, Alex Malbon, Michael Behnke, Rhiannon Jenkinson, I'ah Donovan-Banfield, Stuart Siddell, Kate Heesom and Christy Waterfall, and to the E floor tech team for their endless patience. Thanks to Rayana Kamal and others at the Bristol Animal Rescue Centre, and to all the cats who donated blood and tissue samples; without you, this project would not have been possible.

Last but not least, I dedicate this thesis to my family and other loved ones, human and furry, who bring me joy every day and encourage me to be the best version of myself that I can be.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE:.....

Samantha Saunders

Table of Contents

List of Figures

List of Tables

Abbreviations

A	Adenosine
aa	Amino acid
ADE	Antibody dependent enhancement
APN	Aminopeptidase N
ATP	Adenosine 5' triphosphate
BCoV	Bovine coronavirus
bp	Base pair
C	Cytidine
°C	Degrees Celsius
CCoV	Canine coronavirus
cDNA	Complementary DNA
CER	Cell entry receptor
CO_2	Carbon dioxide
CrFK	Crandell feline kidney cells
C-terminal	Carboxyl-terminal
Da	Dalton
DAPI	4', 6-diamidino-2-phenylindole
dATP	2' deoxyadenosine 5' triphosphate
dCTP	2' deoxycytidine 5' triphosphate
dGTP	2' deoxyguanosine 5' triphosphate
dTTP	2' deoxythymidine 5' triphosphate
DEPC	Diethyl pyrocarbonate
dNTP	Deoxyribonucleotide triphosphate (equimolar mix of dATP, dCTP, dGTP and dTTP)
DMEM	Dulbecco's modified Eagle media
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
E	Envelope (gene/protein)
E. coli	Escherichia coli
EDTA	ethylene diamine tetra-acetic acid (disodium salt)

ELISA	Enzyme-linked immunosorbant assay
ELISpot	Enzyme-linked immunospot
ER	Endoplasmic reticulum
f	Femto
FACS	Fluorescent antibody cell sorting
FBS	Foetal bovine serum
FCoV	Feline coronavirus
FCWF	*Felis catus* whole foetus cells
FECV	Feline enteric coronavirus
FIP	Feline infectious peritonitis
FIPV	Feline infectious peritonitis virus
g	gram
g	gravitational constant
G	Guanosine
GFP	Green fluorescent protein
GTP	Guanosine 5' triphosphate
HCl	Hydrochloric acid
HCoV	Human coronavirus
HE	Haemagglutinin esterase
H_2O	Water
HR1	Heptad repeat 1
HR2	Heptad repeat 2
HRP	Horseradish peroxidase
Hsc70	Heat shock cognate 71 kDa protein
HSP70	Heat shock protein 70 family
HSPA1A	Heat shock 70 kDa protein 1a (a.k.a. Hsp70)
HuNoV	Human norovirus
IBV	Infectious bronchitis virus
ICC	Immunocytochemistry
IFN	Interferon
Ig	Immunoglobulin

IH	Interhelical
IHC	Immunohistochemistry
kb	Kilobase
kbp	Kilobase pair
kDa	Kilodalton
l	litre
LB	Luria Bertani medium
M	Membrane (gene/protein)
MAb	Monoclonal antibody
MEM	Modified Eagles media
MERS	Middle East respiratory syndrome
m	Milli
MHC	Major histocompatibility complex
MHV	Mouse hepatitis virus
M	Molar
mol	Mole
MOPS	3-[N-morpholino] propane sulphonic acid
mRNA	Messenger RNA
N	Nucleocapsid (gene/protein)
n	Nano
nsp	Non-structural protein
nt	Nucleotide
N-terminal	Amino-terminal
ORF	Open reading frame
p	Pico
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PBMC	Peripheral blood mononuclear cells
PCR	Polymerase chain reaction
PEDV	Porcine epidemic diarrhoea virus
PFU	Plaque forming unit

PHA	Phytohaemagglutinin
PMA	Phorbol 12-myristate 13-acetate
Poly(A)	Polyadenylated
pp1a	Polyprotein 1a
pp1ab	Polyprotein 1ab
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT	Reverse transcription
RTC	Replicase-transcriptase complex
RT-PCR	Reverse transcription polymerase chain reaction
S	Spike (gene/protein)
SARS	Severe acute respiratory syndrome
SDS	Sodium dodecyl sulphate
sgRNA	Subgenomic RNA
sgmRNA	Subgenomic messenger RNA
SGTA	Small glutamine-rich tetratricopeptide repeat-containing protein alpha
ssRNA	Single stranded RNA
T	Thymine
Taq	Heat-stable DNA polymerase isolated from *Thermus aquaticus*
TBE	Tris borate EDTA
TGEV	Transmissible gastroenteritis virus
Th1	T helper 1
TMT	Tandem mass tag
Tris	Tris (hydroxymethyl) methylamine
U	Uracil
UTR	Untranslated region
UV	Ultraviolet
µ	Micro
v/v	Volume per volume
WFE	Whole feline embryo cells
w/v	Weight per volume

1 Introduction

1.1 Coronaviruses

The *Coronavirinae* are a sub-family of viruses belonging to the family *Coronaviridae* and order *Nidovirales*. They are spherical to pleomorphic, enveloped, positive sense ssRNA viruses with non-segmented genomes of around 30,000 nucleotides (Siddell, 1995). Coronaviruses exhibit a high rate of mutation during RNA replication and therefore exist as quasispecies, or clusters of genetically diverse populations. This genetic diversity, along with a readiness to recombine with other strains, promotes pathogenesis and cross-species transmission (Denison *et al.*, 2011). It has been hypothesised that the sudden emergence of feline coronavirus-related disease in the late 1950s occurred as a result of a jump into cats from another species (Pedersen, 2009).

Coronaviruses can be found in virtually every species of mammal and bird and usually infect the intestinal and/or respiratory tracts (Pedersen, 2014b). Coronaviruses were not considered important pathogens of humans until the emergence in 2002 of severe acute respiratory syndrome coronavirus (SARS-CoV): a coronavirus outbreak that was introduced into the human population, most likely from bats, and resulted in the deaths of around 10% of those infected (Yip *et al.*, 2009). Middle East respiratory syndrome coronavirus (MERS-CoV), a coronavirus thought to have originated from bats but spilt over into the human population *via* dromedary camels as an intermediate host, subsequently caused a SARS-like disease outbreak with a 60% mortality rate (de Groot *et al.*, 2013, Alagaili *et al.*, 2014). Coronaviruses are also clinically and economically important pathogens of a range of domestic species including pigs, chickens, dogs and cats.

1.2 Feline coronavirus

The feline coronavirus (FCoV) sub-species belongs to the *Alphacoronavirus 1* species (Figure 1.1) alongside canine coronavirus (CCoV) and porcine transmissible gastroenteritis virus (TGEV).

Present in most cat populations worldwide, FCoV is a pathogen of domestic, feral and some wild cats (Brown *et al.*, 2009). In most cases, infection with FCoV results in intestinal disease with mild to no clinical signs, but in 5 to 12% of infected cats a fatal syndrome known as feline infectious peritonitis (FIP) develops (Addie and Jarrett, 1992, Addie *et al.*, 1995).

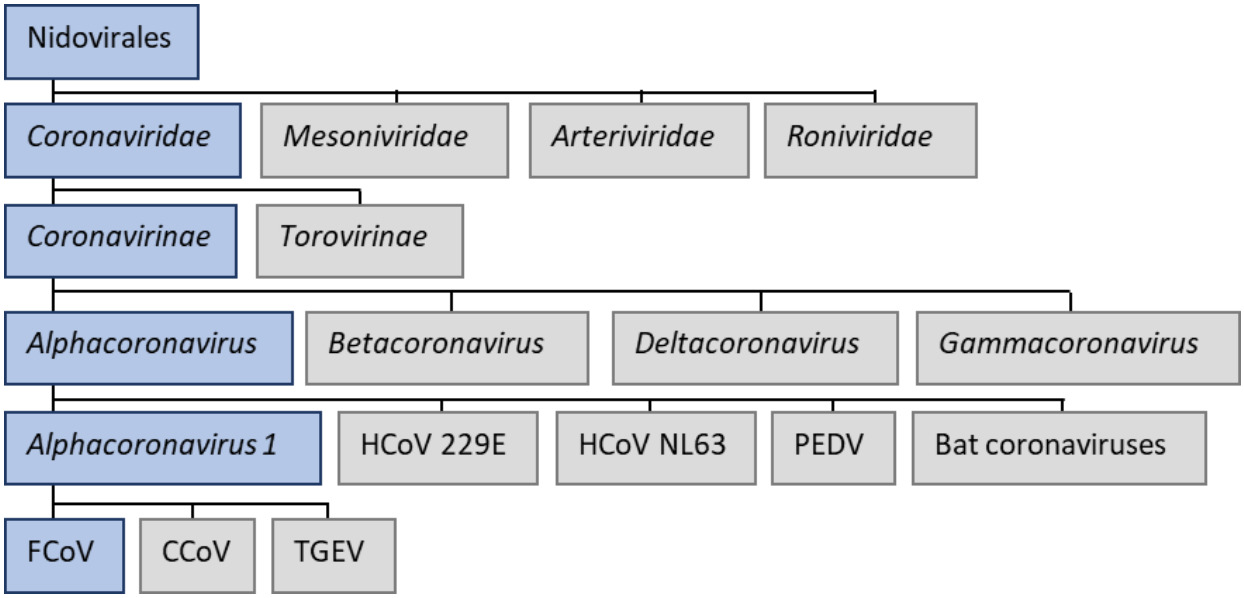

Figure 1.1. Flowchart to show FCoV classification. HCoV: human coronavirus, PEDV: porcine epidemic diarrhoea virus; adapted from Kipar and Meli (2014).

1.2.1 Serotypes of FCoV

Two FCoV serotypes are recognised: Type 1, which represents the vast majority of field strains (Benetka et al., 2004, Addie et al., 2003, Li et al., 2018), and Type 2, which came about through homologous recombination between Type 1 and CCoV. The spike (S) gene is the most significant region of variation between the two serotypes of FCoV; the Type 2 FCoV S gene has a homology of around 91% with that of CCoV, compared to 46% with that of Type 1 FCoV (Jaimes and Whittaker, 2018). Different strains of Type 2 FCoV appear to have arisen through independent recombination events and, depending on the strain in question, further regions originating from CCoV may include the open reading frame (ORF)1b, 3a-c, envelope (E) and membrane (M) genes (Herrewegh et al., 1998).

Distinction between the FCoV serotypes is made using reverse transcription polymerase chain reaction (RT-PCR) to detect characteristic differences in the viral S gene (Addie *et al.*, 2003), or by testing antibodies raised in the infected cat for their ability to detect each serotype (Kummrow *et al.*, 2005, Shiba *et al.*, 2007). Using these methods, studies looking at the relative prevalances of the two FCoV serotypes in Europe and Asia have found that Type 2 FCoV represents 2-11% of natural infections (Addie *et al.*, 2003, Kummrow *et al.*, 2005, Li *et al.*, 2018, Lin *et al.*, 2009, Hohdatsu *et al.*, 1992, Duarte *et al.*, 2009, Shiba *et al.*, 2007).

The S gene encodes the S protein, which mediates host cell entry. Type 2 FCoV uses the cell entry receptor (CER) feline aminopeptidase N (fAPN) (Belouzard et al., 2012): a 150 kDa cell surface glycoprotein commonly expressed in progenitor cells of the granulocyte-monocyte lineage as well as respiratory and intestinal epithelial cells (Hohdatsu et al., 1998). The CER for Type 1 FCoV is

unknown (Dye et al., 2007, Hohdatsu et al., 1998). Both Type 1 and 2 FCoV are capable of causing FIP (Benetka *et al.*, 2004). Two Asian studies found that Type 2 FCoV was overrepresented in cats with FIP (Hohdatsu *et al.*, 1992, Lin *et al.*, 2009), whereas a study conducted in Europe found the converse (Kummrow *et al.*, 2005).

1.2.2 FCoV genome and structure

1.2.2.1 FCoV genome

The FCoV genome (Figure 1.2) comprises two open reading frames (ORF1a and ORF 1b) that encompass around two thirds of the genome and encode 16 non-structural proteins mainly involved in viral RNA synthesis, and nine open reading frames that encode four structural proteins (S, E, M and nucleocapsid (N); Figure 1.3) and five group-specific accessory proteins (3a, 3b, 3c, 7a and 7b) (Kipar and Meli, 2014).

There is an untranslated region at the 5′ end of the coronavirus genome comprising a 65-100 nucleotide leader sequence (Pasternak *et al.*, 2006) followed by a region that is involved in RNA replication (Raman and Brian, 2005), then a transcription regulating sequence (TRS) that is also found preceding each structural and accessory gene and is involved in discontinuous transcription of subgenomic viral RNAs (Dufour *et al.*, 2011). A second untranslated region at the 3′ end of the genome is thought to act as a switch regulating viral RNA synthesis (Goebel *et al.*, 2004).

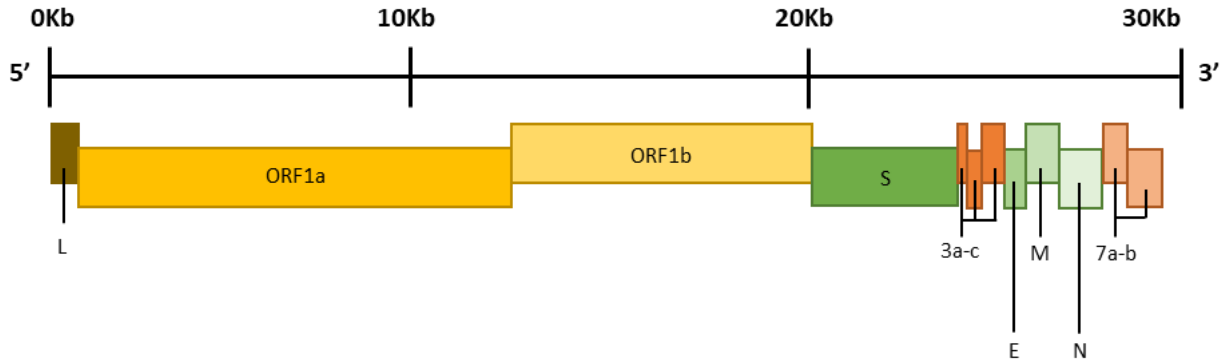

Figure 1.2. Schematic diagram of the FCoV genome. The ~29 kb genome consists of 11 open reading frames (ORFs) encoding 16 non-structural proteins (ORF1a-b), four structural proteins (spike (S), envelope (E), membrane (M), nucleocapsid (N)) and five group-specific accessory proteins (3a-c, 7a-b). L: leader sequence; adapted from Kipar and Meli (2014).

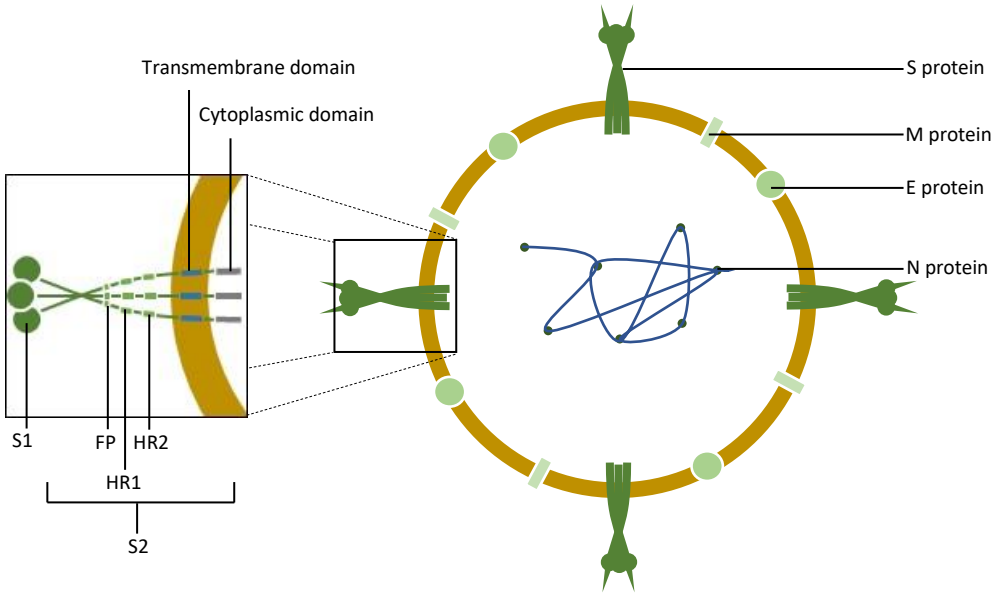

Figure 1.3. Schematic diagram of an FCoV viral particle. The four structural proteins (S: spike, E: envelope, M: membrane, N: nucleocapsid) are displayed. A closer look at the S protein reveals its two functional domains (S1 and S2) and, within S2, the fusion peptide (FP), heptad repeat 1 (HR1) and heptad repeat 2 (HR2) regions. Adapted from Kipar and Meli (2014) and Fields *et al.* (2013).

1.2.2.2 S protein

The S protein is a 180-220 kDa transmembrane protein that mediates host cell entry. It is highly glycosylated and assembles as trimers on the viral particle surface (Jaimes and Whittaker, 2018), giving the virus its characteristic 'crown-like' appearance (Figure 1.4). The S protein consists of three domains, from N-terminus to C-terminus: a large external domain, a transmembrane domain and a short cytoplasmic domain (Fields *et al.*, 2013). Alternatively, the S protein can be divided into S1, which contains the receptor binding domain and is responsible for binding to the target receptor, and S2. S1 comprises an C-terminal domain which binds proteins and a N-terminal domain which is generally responsible for binding carbohydrates (Li, 2016). S2 contains all the features of a class I fusion protein, including a fusion peptide which enables fusion of the virus envelope with the host cell membrane, and two heptad repeats (HR1 and HR2), separated by a stretch of ~140 amino acids called the interhelical (IH) region, which form coiled coils that participate in the fusion process (Bosch *et al.*, 2003, Jaimes and Whittaker, 2018, Belouzard *et al.*, 2012). The S2 region of both FCoV serotypes contains a cleavage site (S2'), while a cleavage site at the border of S1 and S2 is only present in Type 1 FCoV (Jaimes and Whittaker, 2018). Through mediating receptor binding and membrane fusion, the S protein determines the tropism of the virus (Belouzard *et al.*, 2012).

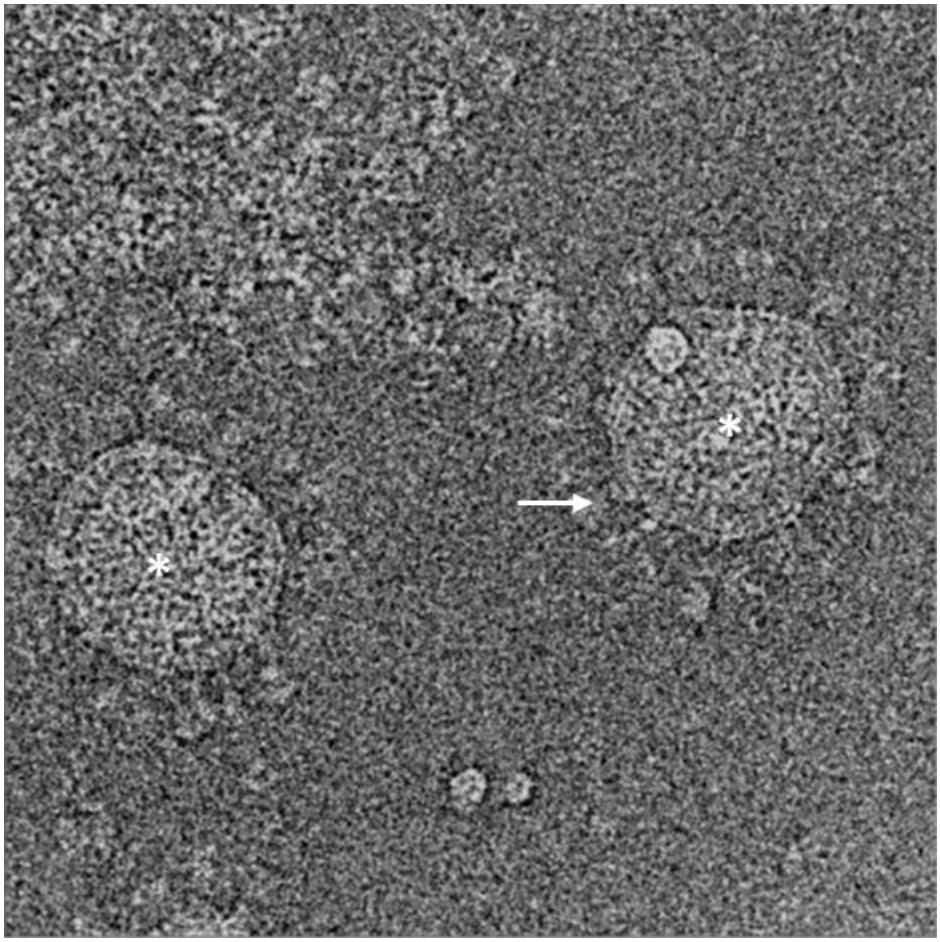

Figure 1.4. Cryo-electron micrograph of two FCoV particles (marked with white asterisks). This image illustrates the spherical shape and distinctive S protein projections (one of which is indicated by a white arrow) of the viral particles, giving a 'crown-like' appearance. Reproduced with permission from Neuman *et al.* (2011).

The cryo-electron microscopy structure of the FCoV S protein has not been reported, but Wu *et al.* (2009) used this method to solve the structure of the S protein of another *Alphacoronavirus*, human coronavirus (HCoV) NL63, complexed with its CER, angiotensin-converting enzyme 2 (ACE2). The study found that the HCoV NL63 receptor binding domain consisted of a novel β-sandwich core structure comprising two layers of β-sheets, which presented three receptor binding motifs to bind ACE2 (Wu *et al.*, 2009). The receptor binding domain of porcine respiratory coronavirus, an *Alphacoronavirus* closely related to TGEV that uses porcine APN as its CER, has a very similar core structure to HCoV NL63 but different receptor binding motifs, which account for the difference in receptor usage between the two viruses (Li, 2015). Based on the structure of the HCoV NL63 S protein, Jaimes and Whittaker (2018) used *in silico* protein modelling to predict the structure of three strains of FCoV: Type 1 FCoV Black, Type 2 FCoV 79-1146 and Type 2 FCoV 79-1683 (Figure 1.5). The models corroborated the features that had already been identified for other coronavirus S proteins and revealed a conserved organisation for *Alphacoronavirus* S proteins.

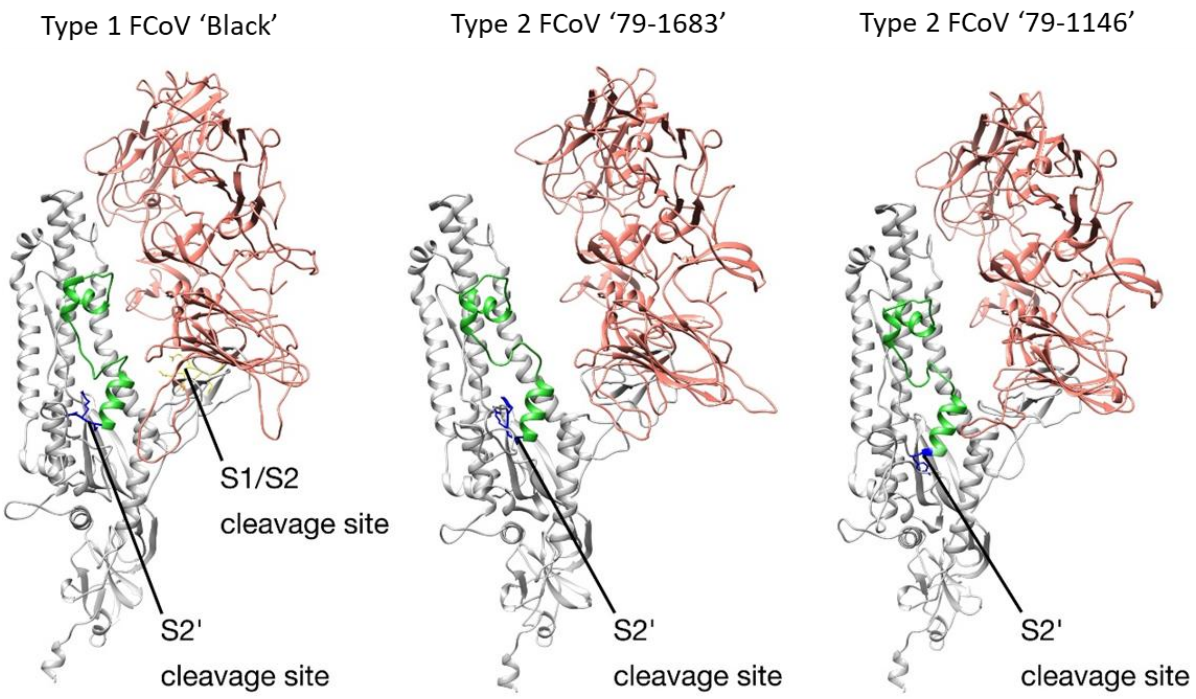

Figure 1.5. In silico protein models of the FCoV S protein. The models are based on the cryo-electron microscopy structure of HCoV NL63 (RCSB PDB # 5SZS). Three strains of FCoV are shown, representing the two serotypes. The S1 domain is shown in salmon, the S2 domain is shown in grey and the fusion peptide is shown in lime green. All three S proteins contain an S2' cleavage site, but only the Type 1 FCoV contains an S1/S2 cleavage site. Reproduced with permission from Jaimes and Whittaker (2018).

1.2.2.3 E protein

The E protein is an 8-12 kDa type III membrane protein, present in the envelope of the viral particle, that has a hydrophilic amino terminus of 7-12 amino acids, a hydrophobic transmembrane domain of 25 amino acids and a hydrophilic carboxyl terminus of 39-77 amino acids. The E protein is localised to the endoplasmic reticulum and intermediate compartment; a complex situated between the endoplasmic reticulum and Golgi apparatus. As well as having ion channel activity, the E protein is known to play a role in virus maturation, assembly, budding and interaction with the host cell (Schoeman and Fielding, 2019).

1.2.2.4 M protein

The 25-30 kDa M protein is the most abundant structural protein in the virus (Fields *et al.*, 2013). It is an N-linked glycosylated class III protein of approximately 230 amino acids, randomly distributed throughout the envelope of the viral particle and comprising a short ectodomain, three transmembrane domains and an extensive endodomain (Jaimes and Whittaker, 2018). The M protein interacts with itself and the other viral structural proteins, and is the main driver of the virus budding process (Alsaadi and Jones, 2019). With the N protein, the M protein forms the structural core of the virion (McBride *et al.*, 2014).

1.2.2.5 N protein

The N protein is a 50 kDa protein that binds and protects viral RNA, forming the viral helical nucleocapsid and enabling transcription. It comprises an N-terminal domain, a C-terminal domain and an intrinsically disordered central region, which all bind RNA but through different mechanisms. The N protein is localised to the nucleolus and cytoplasm of the host cell (McBride *et al.*, 2014). Phosphorylation of the N protein may be responsible for its bias towards binding viral rather than non-viral RNA (Jaimes and Whittaker, 2018). The N protein plays an important role in viral RNA replication and transcription by tethering RNA to the replicase-transcriptase complex (RTC; section 1.2.2.6), and as such is required for efficient viral RNA synthesis (Almazan *et al.*, 2004, Hurst *et al.*, 2010).

1.2.2.6 Non-structural proteins

The coronavirus RTC is encoded by ORF1a and ORF1b. A heptanucleotide 'slippery' sequence and an RNA pseudoknot sit between these ORFs (Namy *et al.*, 2006). In most cases, the ribosome unwinds the pseudoknot and ORF1a is translated into polyprotein 1a (pp1a). However, in some cases the ribosome is blocked by the pseudoknot and becomes stuck on the slippery sequence, which causes the ribosome to shift back by one nucleotide and continue translation using a different reading frame. This results in the longer pp1ab. The two polyproteins are processed into 11 and 16 active subunits respectively by viral-encoded proteases (Fehr and Perlman, 2015). These subunits, also known as 'nsp's, are responsible for replication of the viral genome and generation of template RNAs for transcription of the structural and accessory proteins (Jaimes and Whittaker, 2018).

The functions of the nsps are approximately as follows: nsp1-2 interfere with host defences, nsp3-6 contain factors for the formation of viral replicative organelles and proteinases for the processing of the viral polyproteins, nsp7-11 are involved with viral RNA synthesis through primer-synthesis and interaction with downstream RNA synthesis factors, and nsp12-16 contain enzyme activities necessary for RNA replication, capping and proofreading (Neuman *et al.*, 2014).

1.2.2.7 Accessory proteins

The role of the viral accessory proteins is not completely understood. ORF3a-c are well conserved within the *Alphacoronavirus 1* species. ORF3a and ORF3b encode proteins that are localised in the nucleus and cytoplasm, and the mitochondrion and nucleolus respectively, and whose functions are unknown (Meszaros *et al.*, 2018). The predicted sequence of ORF3c suggests that it encodes a membrane-residing protein with a similar structure to M protein (Kipar and Meli, 2014).

ORF7a encodes a small membrane protein that acts as an interferon-alpha (IFN-α) antagonist (Dedeurwaerder *et al.*, 2014). ORF7b, present only in FCoV and very closely related viruses, encodes a soluble glycoprotein that resides in the lumen of the endoplasmic reticulum (Vennema *et al.*, 1992) and is released into the extracellular space. It is thought that the secreted 7b protein could act as a modulator of the host immune response (Rottier, 1999), and deletions of ORF7b readily occur *in vitro* suggesting that its presence provides a selective advantage *in vivo* (Herrewegh *et al.*, 1995).

1.3 FCoV replication cycle

Figure 1.6 gives an overview of the entire coronavirus replication cycle, then each step is discussed in more detail in sections 1.3.1-4.

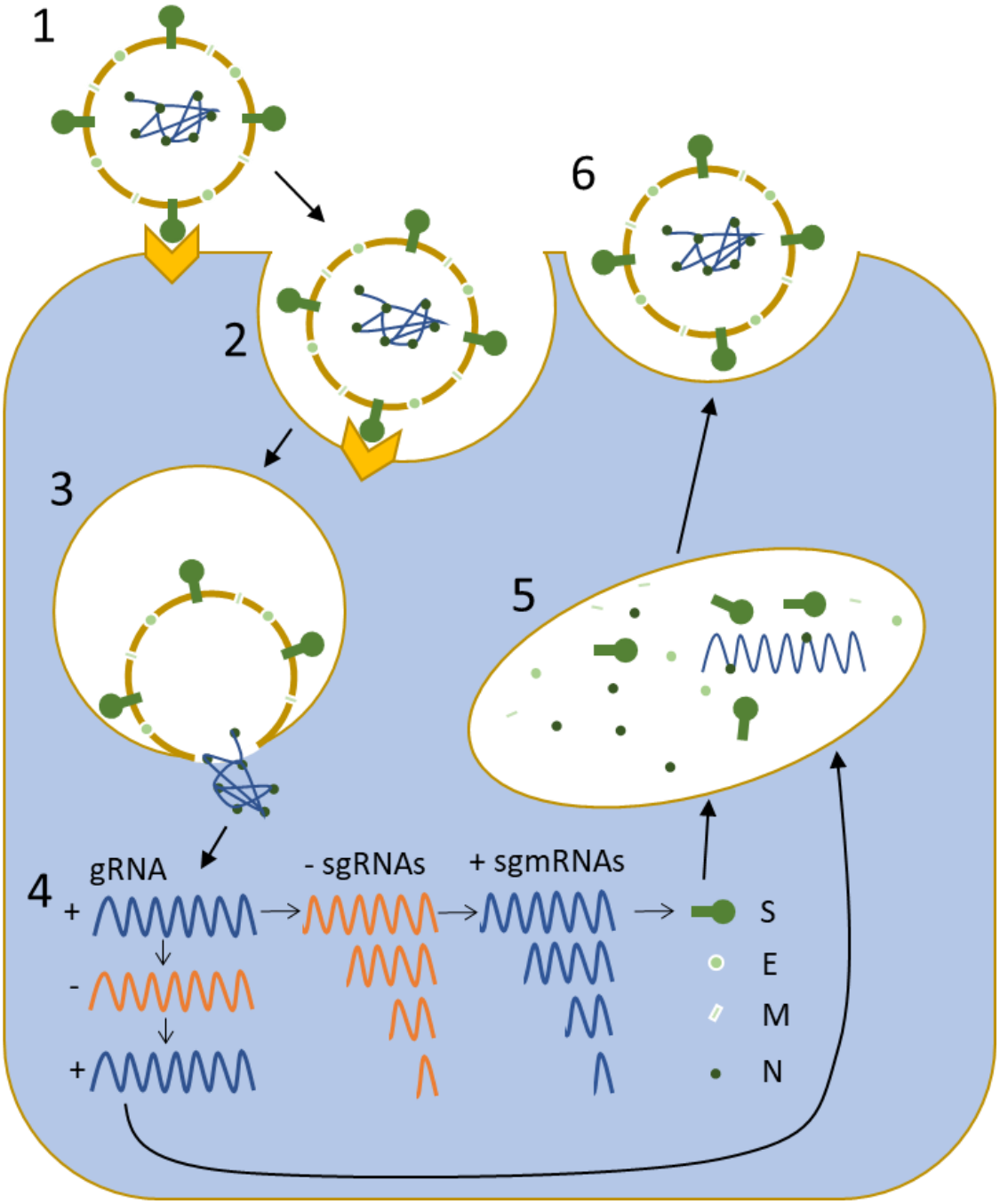

Figure 1.6. A schematic overview of the coronavirus replication cycle. Adapted from Jaimes and Whittaker (2018).
1: The virion recognises and attaches to its receptor(s) on the host cell surface *via* the S protein.
2: The virion is endocytosed.
3: The membrane of the virion fuses with the membrane of the endosome and the viral RNA is released into the cell's cytoplasm.
4: The viral positive-sense (+) genomic RNA (gRNA) undergoes replication to produce negative-sense (-) gRNA, which acts as a template for new copies of the viral genome. The gRNA also undergoes transcription to produce a nested set of negative-sense subgenomic RNAs (sgRNAs), which act as templates for positive-sense messenger sgRNAs (sgmRNAs) from which the structural (S, E, M and N) and accessory (not shown) proteins of the virus are translated.
5: Following maturation in the endoplasmic reticulum (not shown), the structural and accessory proteins, along with newly synthesised copies of the viral genome, assemble in the intermediate compartment.
6: The assembled virions bud from the intermediate compartment and are transported to the cell surface in secretory vesicles, where they are released through fusion of the plasma membrane with the vesicle.

1.3.1 Receptor recognition

The initial step of viral replication is the binding of the viral particle to the plasma membrane of the target cell, which is mediated by the S1 domain of the S protein. The coronavirus S1 domain can be further divided into a C-terminal domain and an N-terminal domain. The C-terminal domain binds protein receptors and the N-terminal domain binds carbohydrate receptors, except in the case of mouse hepatitis virus (MHV) where the N-terminal domain binds a protein receptor (Li, 2015).

Based on the CERs of Type 2 FCoV and other Alphacoronaviruses it is likely that Type 1 FCoV uses a protein receptor. However, the possibility that, like the Betacoronavirus HCoV-OC43 (Szczepanski *et al.*, 2019), Type 1 FCoV utilises a carbohydrate as its CER cannot be ruled out. Additionally many Beta- and Gammacoronaviruses use carbohydrate attachment factors (Li, 2015); receptors that allow the virus to bind to host cell membrane but do not mediate viral entry into the cell (Jolly and Sattentau, 2013). See Table 1 for a summary of the receptors used by different coronaviruses.

A C-type lectin, dendritic cell-specific ICAM-3-grabbing nonintegrin (DC-SIGN), has been shown to play a role in infection of monocytes by Type 2 FCoV and the culture-adapted Type 1 FCoV Black, but is considered more likely to function as an attachment factor than a CER (Regan and Whittaker, 2008, Van Hamme *et al.*, 2011).

Table 1.1. The CERs and attachment factors used by coronaviruses. Where a CER or attachment factor has not been identified, the cell has been left blank. The receptors shown in bold are carbohydrates, while the remainder are proteins. PEDV: porcine epidemic diarrhoea virus, BCoV: bovine coronavirus, IBV: avian infectious bronchitis virus. Adapted from Jaimes and Whittaker (2018), with the source of additional information cited as appropriate.

Genus	Virus	Cell entry receptor	Attachment factor
Alphacoronavirus	TGEV	Porcine/feline aminopeptidase N (Tresnan et al., 1996)	
	Type 1 FCoV		Dendritic cell-specific ICAM-3-grabbing nonintegrin (Regan et al., 2010)
	Type 2 FCoV	Feline/canine aminopeptidase N (Benbacer et al., 1997)	Dendritic cell-specific ICAM-3-grabbing nonintegrin (Regan et al., 2010)
	Type 2 CCoV	Canine/feline aminopeptidase N (Licitra et al., 2014, Tresnan et al., 1996)	
	PEDV	Porcine aminopeptidase N	
	HCoV-NL63	Angiotensin converting enzyme 2	
	HCoV-229E	Human/feline aminopeptidase N (Tresnan et al., 1996)	
Betacoronavirus	BCoV	Major histocompatibility complex class I (Szczepanski et al., 2019)	**9-*O*-acetylated neuraminic acid** (Szczepanski et al., 2019)
	HCoV-OC43	**9-*O*-acetylated neuraminic acid** (Szczepanski et al., 2019)	
	MERS-CoV	Dipeptidyl peptidase 4	**α-2,3-linked sialic acid** (Widagdo et al., 2019)
	MHV	Murine carcinoembryonic antigen-related adhesion molecule 1	
	SARS-CoV	Angiotensin converting enzyme 2	
Gammacoronavirus	IBV		**α-2,3-linked sialic acid** (Cavanagh, 2007), Heat shock protein 70 (Zhang et al., 2017)
	IBV Beaudette		**Heparan sulfate** (Madu et al. 2007)

1.3.2 Cell entry

Following receptor binding, the virus must fuse with a cell membrane in order to gain entry to the cell. This can occur either directly at the cell surface or in the endosomal compartment following endocytosis (Belouzard *et al.*, 2012). Coronaviruses in general can use either mechanism, but FCoV has been demonstrated to use the endocytic pathway *in vitro* (Van Hamme *et al.*, 2007). Once endocytosed, it is thought that the Type 2 FCoV S protein is primed for membrane fusion through cleavage by the cysteine protease cathepsin and a drop in endosomal pH (Regan *et al.*, 2008). Less is known about the mechanism of entry for Type 1 FCoV, but a functional S protein furin or heparan sulfate cleavage site at the boundary of S1 and S2 identified in some strains suggests that Type 1 FCoV S protein likely undergoes cleavage prior to fusion (Licitra *et al.*, 2013, de Haan *et al.*, 2008). S protein cleavage is thought to activate the fusion function of the S2 subunit (Burkard *et al.*, 2014), allowing it to fuse with the endosomal membrane and the viral RNA to be released into the cytoplasm of the host cell (Pedersen, 2014b).

Anti-coronavirus antibodies have been demonstrated to facilitate uptake of virus into macrophages *via* Fc receptors *in vitro*, bypassing the requirement for receptor binding (Takano *et al.*, 2008a). Additionally, transcriptional profiling of peritoneal cells showed upregulation of Fc receptor genes in cats with FCoV-related disease (Watanabe *et al.*, 2018), suggesting that this mechanism may also occur *in vivo*. Increased uptake of virus by macrophages is thought to manifest clinically in antibody dependent enhancement of disease, whereby cats immunised against FCoV prior to experimental infection show decreased survival rates and times post-infection (Takano *et al.*, 2008b, Klepfer *et al.*, 1995, Vennema *et al.*, 1990, Balint *et al.*, 2014).

1.3.3 Viral RNA replication, transcription and translation

ORF1a and ORF1b are immediately translated following virus uncoating, which results in the production of pp1a and pp1ab. These polyproteins are cleaved into nsp1-16, many of which go on to form the RTC (Fehr and Perlman, 2015). The RTC recognises *cis*-acting elements at the 5′ and 3′ untranslated regions of the viral genome and copies the genome either continuously for replication, or discontinuously for transcription (Sawicki *et al.*, 2007). In the case of replication, the full-length negative-sense RNA generated is used as a template for making positive-sense copies of the genome. The N protein participates in viral RNA synthesis by tethering the RNA to the RTC (Hurst *et al.*, 2010).

Discontinuous transcription generates a nested set of seven negative-sense subgenomic RNAs (sgRNAs) which act as templates for positive-sense subgenomic mRNAs (sgmRNAs), from which the

structural and accessory proteins of the virus are translated (Figure 1.7). Transcription proceeds from the 3′ end of the genome and, when a TRS is reached, either continues or stops and jumps to the leader sequence. In this way each sgRNA contains a common leader sequence derived from the 5′ end of the viral genome (Sola *et al.*, 2015, Sawicki *et al.*, 2007). This strand switching capability of the RNA-dependent RNA polymerase is what allows coronaviruses to undergo homologous recombination (Fehr and Perlman, 2015).

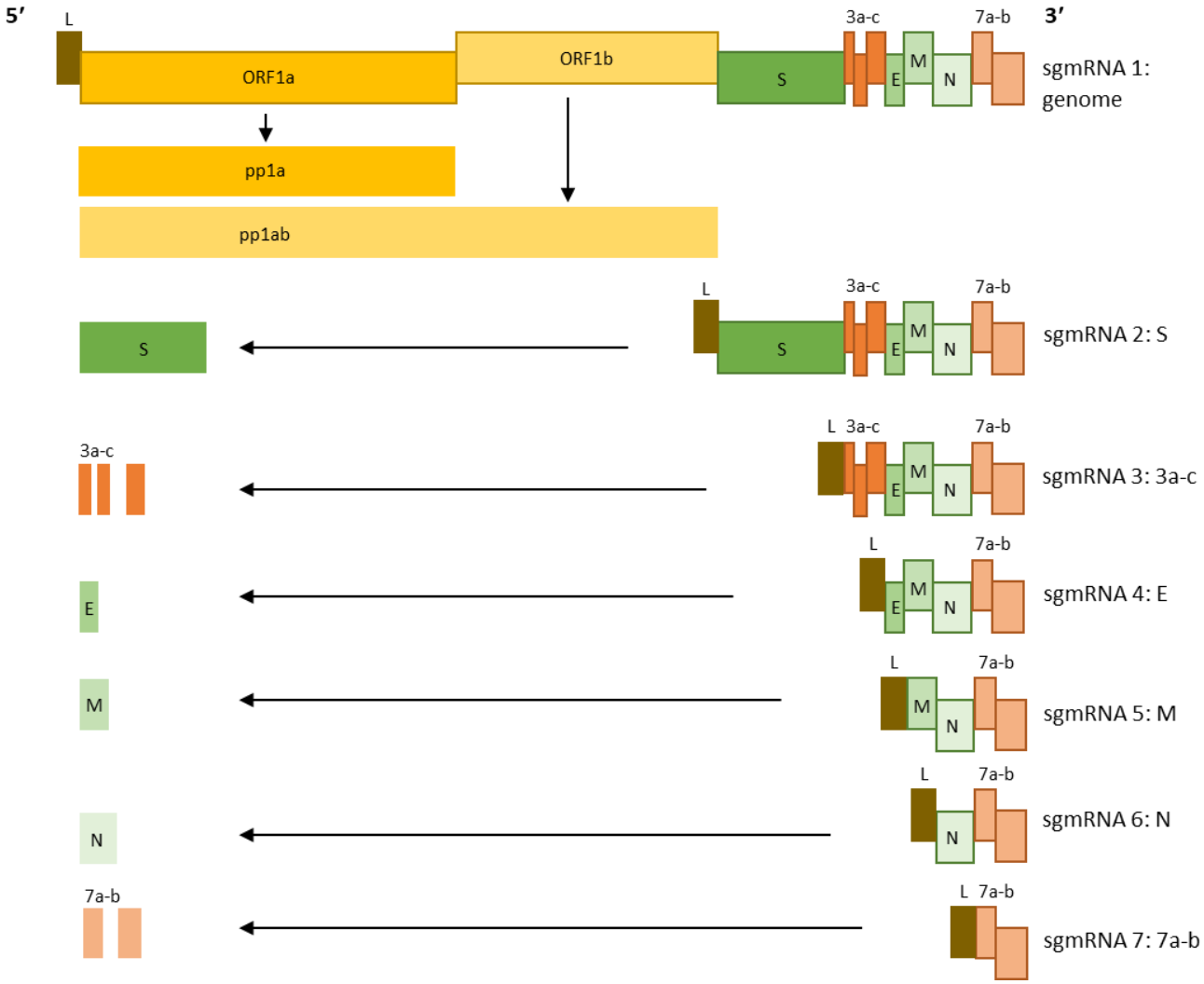

Figure 1.7. Schematic to show the transcription and translation of FCoV sgRNAs. The viral genome is transcribed discontinuously from the 3′ end to produce a set of seven nested negative-sense sgRNAs (not shown), which act as templates for the positive-sense sgmRNAs. The full-length genome (sgmRNA 1) acts as the first sgmRNA, from which pp1a and pp1ab are translated. The structural (S, E, M, N) and accessory (3a-c, 7a-b) proteins are translated from the subsequent sgmRNAs. L: leader sequence. Adapted from Sawicki *et al.* (2007).

1.3.4 Viral particle assembly and budding

The viral structural proteins mature in the endoplasmic reticulum and are transported through the secretory pathway into the intermediate compartment, which is the site of coronavirus assembly. The M protein recruits a region of membrane for viral particle assembly and guides other structural proteins to the site, while the N protein binds to the newly synthesised copies of the viral genome to form the nucleocapsid (Neuman *et al.*, 2011). The assembled virions bud from the intermediate

compartment and are transported in secretory vesicles to the cell surface and released through endosome-plasma membrane fusion (Rottier, 1999, Jaimes and Whittaker, 2018).

1.4 Clinical syndromes caused by FCoV

FCoV infection causes two distinct disease types in cats: a transient to persistent enteric disease that usually involves mild clinical signs if any, and a near invariably fatal systemic disease known as FIP. The virus pathotype that results in the former clinical presentation is referred to as 'feline enteric coronavirus' (FECV), while the pathotype that results in the latter is referred to as 'FIP virus' (FIPV) (Pedersen, 2009). It is thought that, in most cases, a cat will initially become infected with FECV, then the virus will mutate within that cat to FIPV (Pedersen, 2014b). Both Type 1 and 2 FCoV can be either FECV or FIPV (Benetka *et al.*, 2004).

1.4.1 FECV

FECV is present in most cat populations and extremely contagious, with a seroprevalence of 4 to 24% in single cat households. This figure rises to between 26 and 87% in multi-cat environments (Drechsler *et al.*, 2011), and FCoV was found to be by far the most common faecal pathogen among cats living in a shelter in the USA (Sabshin *et al.*, 2012). This demonstrates the ubiquity of FCoV, especially amongst cats living in large groups such as those in shelters or catteries. Typical clinical signs of FECV include diarrhoea and inappetence, though many infected cats will have no signs at all (Pedersen *et al.*, 2008). A very small minority of cats die from FECV-associated enteritis (Kipar *et al.*, 1998b). Transmission of FECV occurs *via* the faecal-oral route (Meli *et al.*, 2004), and the virus infects the mature epithelial cells of the small and large intestine (Kipar *et al.*, 1998b).

Following infection, cats start shedding FECV in their faeces within a week and continue to shed for up to 11 months. They then either clear the virus and cease shedding (though remain susceptible to reinfection with the same or a different strain) or continue to shed persistently, possibly for life. Persistent infection does not appear to predispose a cat to development of FIP (Pedersen *et al.*, 2008, Addie *et al.*, 2003, Addie and Jarrett, 2001). Approximately 3% of cats are naturally resistant to infection with FECV (Addie and Jarrett, 2001).

It was previously thought that, unlike FIPV, FECV is confined to the gut and unable to infect monocytes and macrophages (Pedersen, 1987). However, it has been demonstrated through detection of FCoV in the blood of clinically healthy cats that FECV can infect monocytes and spread from the gut *via* monocyte-associated viraemia (Kipar *et al.*, 1999a, Gunn-Moore *et al.*, 1998, Meli *et al.*, 2004, Desmarets *et al.*, 2016). It was demonstrated in one case that FECV can infect monocytes even without an initial period of intestinal replication (Desmarets *et al.*, 2016). Gunn-Moore *et al.*

(1998) found that 80% of FCoV-infected cats without FIP in endemic households had FCoV RNA in circulating monocytes. However, the viral load in cats without FIP is significantly lower than those with FIP (Kipar *et al.*, 2006), perhaps indicating that it is not systemic infection but the ability of the virus to replicate to high levels in monocytes that is the key event leading to FIP. This has been demonstrated in feline monocytes infected *in vitro* with FCoV, whereby FIPV can replicate sustainably and to a much higher level than FECV in this cell type (Dewerchin *et al.*, 2005). Once systemic, the virus can reside within macrophages in extra-intestinal sites including the abdominal organs, lymphoid tissues and brain, still without resulting in any clinical signs. The virus can persist in these sites long after viraemia has cleared. The major site of FECV persistence in the gut is the colon (Kipar *et al.*, 2010).

FCoV-infected cats that do not go on to develop FIP demonstrate a strong immune response to the virus, including rising antibody titre, circulating immune complexes, and follicular and T cell hyperplasia in the lymph nodes, thymus and spleen (Kipar *et al.*, 1999a). These tissues contain FCoV-specific plasma cells, indicating a specific response to FCoV rather than a general response to inflammation (Kipar *et al.*, 1998a).

1.4.2 FIPV

In 5 to 12% of FCoV-infected cats, FIP develops (Addie and Jarrett, 1992, Addie *et al.*, 1995). This is a usually fatal systemic disease caused by FIPV, which in most cases is thought to be a pathogenic mutant of FECV that has arisen within the individual cat (Pedersen, 2014b). Though the ability to infect monocytes and macrophages is not unique to FIPV, the level to which it replicates in these cells sets it apart from FECV (Dewerchin *et al.*, 2005); substantial replication of FIPV in monocytes results in their activation and precipitates the development of FIP (Kipar and Meli, 2014). When cats are experimentally infected with FIPV, clinical signs of FIP emerge from around 10 days post infection (de Groot-Mijnes *et al.*, 2005). This does not reflect the natural situation, where the incubation period is usually weeks to months (Addie *et al.*, 1995).

Unlike FECV, FIPV is not usually transmitted horizontally. This is because FIPV is rarely shed in the faeces and, when it is, it is not infectious to other cats (Pedersen *et al.*, 2012). Despite this, FIP outbreaks have been reported in multi-cat environments. Host and environmental factors are likely important factors in these outbreaks, but horizontal transmission of FIPV (or at least FCoV with an increased tendency to mutate into FIPV) is also thought to play a part (Drechsler *et al.*, 2011, Kipar and Meli, 2014).

1.4.2.1 Characteristic lesions of FIP

It was previously thought that the characteristic pyogranulomatous perivascular lesions of FIP (Figure 1.8) are the result of immune complex formation and deposition in vessels (Olsen, 1993). However, circulating immune complexes have been found in cats infected with FCoV that do not go on to develop FIP (Kipar *et al.*, 1999a). More recent evidence demonstrates that the lesions of naturally infected cats are monocyte mediated. Activated, FIPV-infected monocytes produce cytokines and adhesion molecules that allow them to interact with activated endothelial cells and extravasate (Acar *et al.*, 2016, Olyslaegers *et al.*, 2013). The activated monocytes also produce enzymes that dissolve the vascular basement membrane (Kipar *et al.*, 2005) and neutrophil survival factors that attract neutrophils to the lesions and give them their pyogranulomatous character (Takano *et al.*, 2009). Lesions are typically restricted to veins, likely due to selective responsiveness of the endothelium (Kipar *et al.*, 2005).

Figure 1.8. Characteristic lesions of feline infectious peritonitis on *post mortem* examination. A perivascular pyogranuloma on the small intestinal serosa of a cat with FIP is indicated with a white arrow. Image courtesy of The Feline Centre, Langford Vets, University of Bristol.

1.4.2.2 Clinical signs of FIP

Most clinical signs of FIP are related to the pyogranulomatous perivascular lesions that characterise the disease. As a result, they can vary depending on which organs are affected. Commonly involved sites are the serosa, kidneys, mesenteric lymph nodes, brain, eyes, liver, spleen and lungs (Kipar and Meli, 2014), leading to clinical signs including ascites (Figure 1.9A), dyspnoea, uveitis (Figure 1.9B), ataxia and seizures (Pedersen, 2014a, Riemer *et al.*, 2015). Papular cutaneous lesions have also been described less commonly as a feature of FIP (Declercq *et al.*, 2008, Redford and Al-Dissi, 2019, Osumi *et al.*, 2018). More general findings include persistent pyrexia that is refractory to antibiotics, lethargy, anorexia, icterus and weight loss (Addie *et al.*, 2009)

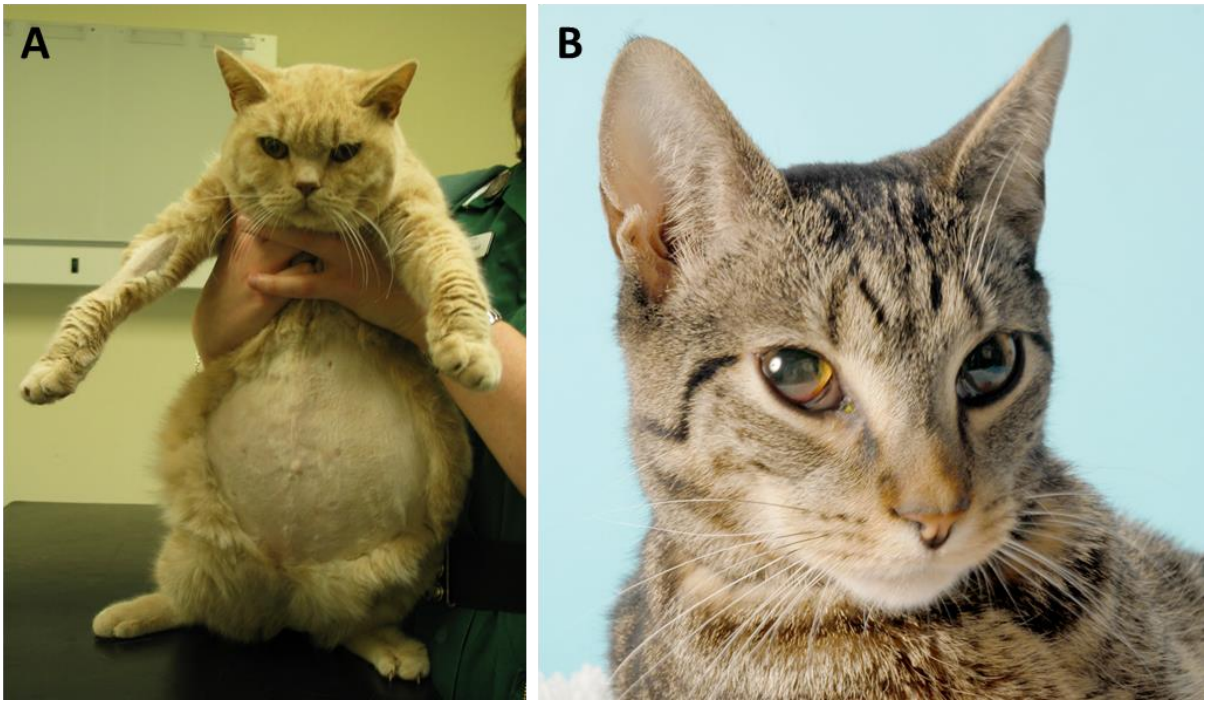

Figure 1.9. Clinical signs in cats with FIP. Image A shows a cat with ascites: accumulation of abdominal fluid due to serosal inflammation. Image B shows a cat with uveitis: inflammation of the eye resulting from pyogranulomatous lesions in the vessels of the region. Images courtesy of The Feline Centre, Langford Vets, University of Bristol.

A distinction is often made between 'wet' and 'dry' FIP. The former presentation involves serosal inflammation and effusions and is associated with a survival time of days to weeks. The latter involves solid parenchymatous lesions in organs and is associated with a survival time of weeks to months (Addie, 2012, Pedersen, 2014a). However, it is more realistic to consider these two presentations the extremes of a spectrum, with most cases sitting somewhere in the middle. This is confirmed by *post mortem* examination of cats with FIP, which usually reveals a combination of effusion and solid lesions (Kipar and Meli, 2014). A form of FIP where lesions appear to be largely restricted to the intestine and/or mesenteric lymph nodes, often leading to a palpable abdominal mass and signs related to intestinal dysfunction (Harvey *et al.*, 1996, Kipar *et al.*, 1999b), could be interpreted as an extreme type of dry FIP.

The prognosis of FIP is extremely poor, with a median survival time from diagnosis of just nine days (Ritz *et al.*, 2007).

1.5 Pathogenesis of FIP

1.5.1 The role of viral mutations in development of FIP

The 'internal mutation' theory of FIP pathogenesis is commonly accepted, whereby the host becomes infected with FECV and the virus acquires mutations within the individual to become FIPV (Pedersen, 2014b). Various viral mutations have been identified that correlate with pathotype, but to date none that definitively distinguish FIPV from FECV have been found. Differences have been noted between FECVs and FIPVs in the M and ORF7a-b genes (Brown *et al.*, 2009, Kennedy *et al.*, 2001, Herrewegh *et al.*, 1995, Borschensky and Reinacher, 2014), but two regions of particular interest are ORF3c and S.

The ORF3c gene was found to be intact in all FCoVs derived from the faeces of healthy cats (presumed to be FECVs) but remained intact in only a small proportion of FCoVs derived from the tissues of cats with FIP (presumed to be FIPVs) (Pedersen *et al.*, 2012, Chang *et al.*, 2010, Bank-Wolf *et al.*, 2014, Borschensky and Reinacher, 2014, Hora *et al.*, 2016, Vennema *et al.*, 1998). This could be interpreted as ORF3c mutations conferring virulence to FCoV, but Chang *et al.* (2010) argued that 3c is only necessary for enteric replication and, once systemic, there is no longer selection pressure on the virus for the gene to remain intact. This was confirmed when FIPVs with intact ORF3c genes were characterised (Pedersen *et al.*, 2012).

Mutations in the S gene have been associated with FCoVs derived from the tissues of cats with FIP but not those derived from the faeces of healthy cats, particularly in the putative S2 fusion peptide (Chang *et al.*, 2012, Lewis *et al.*, 2015, Bank-Wolf *et al.*, 2014), the HR1 region (Bank-Wolf *et al.*, 2014, Lewis *et al.*, 2015) and the cleavage site between S1 and S2 (Licitra *et al.*, 2013). However, the majority of FCoVs in the tissues of healthy cats with systemic FCoV also carry the S2 fusion peptide mutations, suggesting that these mutations are associated with a switch in tropism and systemic spread rather than virulence (Barker *et al.*, 2017, Porter *et al.*, 2014). The relevance of the HR1 mutations to virus pathotype are unknown, though it has been demonstrated that they sit within a major T cell epitope for Type 1 FCoV (Satoh *et al.*, 2011a) so may alter the immune response of the host to the virus. The mutations in the cleavage site between S1 and S2 have been demonstrated to alter the protease activation requirements of the virus (Licitra *et al.*, 2013), offering a possible explanation for the increased ability of FIPVs to replicate in monocytes and macrophages.

By constructing chimeric viruses bearing regions of the S gene from either an FECV or FIPV, Rottier *et al.* (2005) demonstrated that the S2 region dictates macrophage tropism. To investigate this further, Shirato *et al.* (2018) mapped the exact mutations in the Type 2 FIPV S gene that confer its ability to replicate effectively in *ex vivo* feline macrophages. The S genes of FECVs and FIPVs were compared,

and the non-synonymous mutations consistently found in FIPVs only were engineered into a recombinant FECV, using reverse genetics, to observe their effect on virus phenotype. The authors found that efficient replication in macrophages could be pinpointed to five cooperative mutations in the FCoV S2 region. Though these mutations are not identical to those that have previously been identified through comparing the sequences of FECVs and FIPVs, one mutation sits within the S2 fusion peptide and another lies within the HR1 region.

Though viral mutations that enable efficient replication in monocytes and macrophages are undoubtedly necessary for the development of FIP, no mutation or set of mutations is unique to FIPV or sufficient to cause FIP in all cats. When challenged with the same dose of genetically identical FIPV, some cats succumb while others are able to overcome the infection (de Groot-Mijnes *et al.*, 2005, Pedersen *et al.*, 2014, Mustaffa-Kamal *et al.*, 2019). This demonstrates the importance of host factors in conferring resistance and susceptibility to FIP.

1.5.2 The role of the host in development of FIP

Cats with FIP are typically under two years of age, though cats of any age can be affected and a smaller peak is observed in cats older than 10 years (Riemer *et al.*, 2015, Addie and Jarrett, 1992, Addie *et al.*, 1995, Pesteanu-Somogyi *et al.*, 2006). A longitudinal study of cats naturally infected with FCoV found that the risk of developing FIP decreased over time (Addie *et al.*, 1995). This phenomenon could be explained by the development of immunological resistance to FIP but would also be observed if some cats were intrinsically susceptible to FIP and succumbed quickly, leaving those that were intrinsically resistant behind.

Males are slightly more susceptible to FIP than females (Riemer *et al.*, 2015, Rohrbach *et al.*, 2001, Norris *et al.*, 2005), and some studies have found that unneutered males are overrepresented compared to their neutered and female counterparts (Rohrbach *et al.*, 2001, Pesteanu-Somogyi *et al.*, 2006). However, cats from breeding households are more likely to be unneutered, and it is known that FCoV is more prevalent in multi-cat environments such as breeding households and shelters.

Pedigree cats are more likely to have FIP than non-pedigree cats (Norris *et al.*, 2005, Pesteanu-Somogyi *et al.*, 2006, Rohrbach *et al.*, 2001), though again cats from breeding households are more likely to be pedigree so may be overrepresented. The two major studies into the incidence of FIP in individual breeds disagree over which are susceptible and which are resistant (Pesteanu-Somogyi *et al.*, 2006, Norris *et al.*, 2005). Pedersen (2009) argues that perhaps the bloodlines within a breed are more relevant than the breeds themselves. However, an attempt to breed FIP-resistant cats through

positive genetic selection led to an increase in susceptibility, suggesting that genetic diversity is more important than selection of particular traits (Pedersen *et al.*, 2016).

Based on the idea that gut microbiota may influence the immune response to viral disease, a pilot study identified differences in the intestinal microbial communities between cats with FIP, healthy FCoV-infected cats and cats without FCoV. However, the authors were unable to draw any conclusions about the relevance of these differences (Meazzi *et al.*, 2019).

1.5.2.1 Immune Response to FCoV

It is generally accepted that the pathology associated with FIP arises due to an interplay between the virus and the host's immune system, with the immune response of the individual guiding disease outcome (Pedersen, 2014b). A study following a cohort of cats experimentally infected with FIPV found that antibody titre rises at a similar rate and magnitude in all cats, regardless of clinical course or outcome (de Groot-Mijnes *et al.*, 2005), demonstrating that a humoral response to FIPV infection takes place but is not protective. This is also evidenced by the presence of B cells with anti-FCoV antibodies in the lymphoid tissues of both cats with FIP and those that overcame the infection (Kipar *et al.*, 1999a) and the high antibody titres often found in cats with naturally occurring FIP (Hartmann *et al.*, 2003). Cornelissen *et al.* (2009) demonstrated that FIPV-infected macrophages do not undergo antibody-dependent, complement-mediated lysis, suggesting that the usual protective function of antibodies is somehow subverted in FIP.

Antibody dependent enhancement (ADE) has been documented in experimentally infected cats whereby passive or active immunisation, particularly against the FIPV S protein, actually decreases survival rate and survival time post-infection (Takano *et al.*, 2008b, Klepfer *et al.*, 1995, Vennema *et al.*, 1990, Balint *et al.*, 2014, Takano *et al.*, 2019b). A similar effect has been demonstrated *in vitro*, whereby presence of anti-FIPV antibodies increased uptake of FIPV *via* the macrophage Fc receptor (Corapi *et al.*, 1992, Olsen *et al.*, 1992). However, Addie *et al.* (1995) found that ADE of disease did not occur in naturally infected cats, and that repeated exposure to FCoV did not increase the risk of developing FIP. In cats naturally infected with FCoV, those that cleared the virus had a significantly higher anti-S to anti-M antibody ratio compared with those that shed persistently or succumbed to FIP (Gonon *et al.*, 1999), suggesting that a humoral immune response to the S protein could be protective in at least some situations.

The cellular immune response to FCoV infection seems to be more closely correlated with course and outcome of disease. Following experimental infection with FIPV, all cats initially became T cell depleted. Those cats that succumbed to FIP remained T cell depleted, whereas those that overcame the challenge experienced a strong resurgence in T cell numbers (de Groot-Mijnes *et al.*, 2005,

Mustaffa-Kamal *et al.*, 2019). This reflects the clinical situation, where lymphopenia is a common haematological finding in cats with FIP (Riemer *et al.*, 2015). Studies looking at the lymphoid tissues of cats with FIP and those of FCoV-infected cats without FIP found that the former group had lymphoid depletion and thymic atrophy, compared to the lymphoid hyperplasia seen in the latter group (Kipar *et al.*, 1999a, Kipar *et al.*, 2001). This suggests that a robust cellular immune response may have contained the infection and prevented FIP from developing. *In vitro* studies have demonstrated that T cells, and in particular natural killer cells, from cats with FIP are not only depleted, but are also less functional (Vermeulen *et al.*, 2013, Satoh *et al.*, 2011b). Takano *et al.* (2007) argue that production of tumour necrosis factor alpha (TNF-α) by infected macrophages is responsible for induction of T cell apoptosis, and indeed TNF-α was found to be elevated in the blood (Kiss *et al.*, 2004) and mediastinal lymph node (Malbon *et al.*, 2019) of cats with FIP.

It was previously thought that wet FIP arose when a strong humoral immune response to FCoV infection led to increased uptake of virus into macrophages and immune complex-mediated vasculitis, while dry FIP resulted from a mixed humoral and cellular immune response. The contribution of antibodies to the pathogenesis of the disease cannot be completely ruled out, but it is now understood that lesions in all forms of FIP are primarily mediated by monocytes (Kipar and Meli, 2014).

1.5.3 The role of the environment in development of FIP

It is known that FIP frequently occurs in multi-cat environments such as catteries and shelters (Addie *et al.*, 2009), with the risk for FIP increasing with the number of cats kept together (Kass, 1995), the overall frequency of FCoV shedding and the proportion of cats chronically shedding FCoV (Foley *et al.*, 1997). It is unclear whether the rate of FIP is actually increased in these populations, or whether the increased incidence of FIP simply reflects increased prevalence of FCoV. However, stress (Addie *et al.*, 2009), high viral load (Addie and Jarrett, 1992) and co-infections (Foley *et al.*, 1997) are risk factors for development of FIP from FCoV infection, and these factors are likely to be present where many cats are kept together (Drechsler *et al.*, 2011).

Stress is an important factor in the development of FIP in all FCoV-infected cats, not just those in multi-cat environments. In a retrospective study of 231 cats with confirmed FIP, Riemer *et al.* (2015) found that a specific stressful event such as adoption, surgery or vaccination preceded diagnosis in over half of cats.

1.6 Diagnosis of FIP

The vague and varied clinical signs of FIP can make it a challenging disease to diagnose, particularly when an effusion is not present. The clinician will often use a combination of signalment (the cat's age, sex and breed), history, physical examination findings and multiple indirect and direct tests to come to a diagnosis of FIP. Direct detection of FCoV in tissue or effusion macrophages using immunostaining methods is sufficient for a definitive diagnosis of FIP (section 1.6.2.1), but this is not always possible.

1.6.1 Indirect tests

Indirect tests aim to detect changes associated with FIP rather than the virus itself.

1.6.1.1 Haematology and biochemistry

Certain findings on routine haematology and biochemistry are typical of FIP. These are a mild to moderate non-regenerative anaemia, microcytosis, lymphopenia, neutrophilia with or without left shift, hyperbilirubinaemia, hyperglobulinaemia, hypoalbuminaemia and an albumin to globulin ratio of less than 0.8 (Addie *et al.*, 2009, Riemer *et al.*, 2015). However, it is important to note that these findings are neither pathognomonic for FIP nor present in all cats with FIP. Serum levels of the acute phase protein alpha-1 acid glycoprotein (AGP) greater than 1.5 mg/ml are suggestive of FIP, particularly when other criteria fit the diagnosis, but can be elevated with other inflammatory conditions (Duthie *et al.*, 1997, Paltrinieri *et al.*, 2007, Giori *et al.*, 2011, Stranieri *et al.*, 2018).

1.6.1.2 Effusion analysis

The presence of a characteristic effusion, usually in the thorax or abdomen, is highly suggestive of FIP (Pedersen, 2014a). A typical FIP effusion is yellow, clear and viscous with a high protein content (>35 g/l) and low cellularity (<5 x 10^9 nucleated cells/l). The cells are predominantly macrophages and neutrophils (Tasker, 2018). Albumin, globulin and AGP can also be measured in effusion fluid, with an albumin to globulin ratio of less than 0.4 and an AGP of greater than 1.55 mg/ml being highly consistent with FIP (Tasker, 2018, Hazuchova *et al.*, 2016). Immunostaining can be used to look for FCoV in the effusion fluid macrophages (section 1.6.2.1).

1.6.1.3 Histopathology

Samples of tissues affected by FIP lesions can be collected by biopsy *ante mortem* or on *post mortem* examination. Presence of characteristic histopathological changes are considered reliable for diagnosis of FIP (Tasker, 2018), but immunostaining can be used in addition to look for FCoV in tissue macrophages (section 1.6.2.1).

1.6.2 Direct tests

Direct tests aim to demonstrate the presence of FCoV in clinical samples.

1.6.2.1 Immunostaining

Immunocytochemistry (ICC), used to detect FCoV in macrophages found in effusion fluid (Felten *et al.*, 2016, Litster *et al.*, 2013), cerebrospinal fluid (Ives *et al.*, 2013) and aqueous humour (Felten *et al.*, 2017b), was historically considered to be very specific, but two studies found positive results in samples from cats without FIP (Litster *et al.*, 2013, Felten *et al.*, 2016). Tasker (2018) argued that the false positive results in these studies were likely due to the methodology used or suboptimal storage of cytology slides, suggesting that the specificity of ICC is high if the test is carried out properly. False negative results can occur with ICC due to low cellularity of samples and the inability of the method to detect cells with a low virus burden (Tasker, 2018). Detection of FCoV by immunohistochemistry (IHC) in surgical biopsies or tissue samples obtained at *post mortem* examination has a very high positive predictive value and is considered the gold-standard test for diagnosis of FIP (Kipar and Meli, 2014), but obtaining a biopsy can be challenging due to its invasive nature.

1.6.2.2 Polymerase chain reaction (PCR)

RT-PCR can detect FCoV RNA in various clinical samples. Since FCoV spreads systemically in cats without FIP, its detection in sites beyond the gut is not necessarily indicative of FIP. However, FCoV RNA is more commonly found and at higher levels in tissues from cats with FIP compared to those without (Porter *et al.*, 2014, Barker *et al.*, 2017, Stranieri *et al.*, 2018). The use of RT-PCR on blood samples is limited due to a low specificity and sensitivity for FIP (Meli *et al.*, 2004, Felten *et al.*, 2015, Pedersen *et al.*, 2015, Doenges *et al.*, 2016b), but the test has a very high specificity and reasonable sensitivity when used with effusion fluid and mediastinal lymph node fine needle aspirates from cats with FIP-consistent signs (Longstaff *et al.*, 2015, Doenges *et al.*, 2016b, Dunbar *et al.*, 2018), and cerebrospinal fluid from cats with neurological FIP-consistent signs (Doenges *et al.*, 2016a).

RT-PCR and sequencing have been used in combination to attempt to differentiate FECV from FIPV, based on mutations in the S2 fusion peptide thought to be associated with the latter. The inclusion of a sequencing step was found to slightly increase specificity in some cases compared to RT-PCR alone, but none of the studies demonstrated a definitive distinction between FECV and FIPV (Sangl *et al.*, 2018, Felten *et al.*, 2015, Longstaff *et al.*, 2015, Felten *et al.*, 2017a, Barker *et al.*, 2017). This is unsurprising given that the mutations analysed are not unique to FIPV and rather indicate systemic spread of the virus (Porter *et al.*, 2014, Barker *et al.*, 2017).

1.6.2.3 Serology

Detection of anti-FCoV antibodies in the serum cannot be used to diagnose FIP because positive titres are frequently found in healthy FCoV-infected cats and negative results are possible in cats with FIP (Sparkes *et al.*, 1991, Hartmann *et al.*, 2003). However, a very high anti-FCoV antibody titre was found to correlate with FIP (Foley *et al.*, 1997). Measurement of anti-FCoV antibodies in effusion fluid and cerebrospinal fluid is not recommended because neither test is specific for FIP (Hartmann *et al.*, 2003, Boettcher *et al.*, 2007, Soma *et al.*, 2018).

1.7 Treatment of FIP

Corticosteroids are widely used to suppress the inflammation associated with FIP and increase survival time, but there are no controlled studies proving their efficacy. Other immunosuppressive drugs that have been suggested include chlorambucil, cyclophosphamide and salicylic acid, though evidence to prove their benefit is lacking (Addie *et al.*, 2009). There is anecdotal evidence for the benefit of the methylxanthine derivatives propentofylline and pentoxyfylline, which are thought to ameliorate vasculitis through inhibition of pro-inflammatory cytokines. However, a placebo-controlled double-blind trial investigating the efficacy of propentofylline in 23 cats with naturally occurring FIP found no effect of the drug on survival time or quality of life (Fischer *et al.*, 2011).

Ozagrel hydrochloride is a thromboxane synthetase inhibitor that is hypothesised to treat vasculitis through inhibition of platelet aggregation. Watari *et al.* (1998) describe its successful use in two cats with naturally occurring FIP, though it is difficult to draw any conclusions from such a small case series where the diagnosis of FIP was not confirmed by IHC. Chloroquine inhibits viral replication *in vitro* and has been used to treat immune-mediated inflammatory conditions. A placebo-controlled trial carried out in nine cats experimentally infected with FIPV showed an improved clinical score and a non-significant increase in survival time in the treatment groups compared to control group, but this was accompanied by an increase in liver enzyme alanine aminotransferase suggestive of toxicity (Takano *et al.*, 2013). The antiviral compound ribavirin has been shown to prevent FCoV replication *in vitro*, but kittens treated with the drug had shorter survival times and more severe clinical signs than the control group (Weiss *et al.*, 1993).

Ishida *et al.* (2004) reported a complete or partial remission of FIP in eight out of 12 cats treated with feline interferon-omega (IFN-ω) and corticosteroid. However, the study was uncontrolled and the diagnosis unconfirmed by IHC in those cats that survived. A more recent placebo-controlled double-blind trial carried out in 37 cats showed no benefit of IFN-ω, though the treatment group did include one long-term survivor (Ritz *et al.*, 2007). A study involving 74 experimentally infected cats

showed a modest increase in lifespan with human IFN-α treatment (Weiss *et al.*, 1990), but no benefit to survival was demonstrated and the treatment is untested in naturally infected cats.

Following a report of the remission of dry FIP in two out of three cats treated with polyprenyl immunostimulant (PI) (Legendre and Bartges, 2009), there has been a lot of interest in this agent and related immunostimulatory substances. A follow up study saw 60 cats with naturally occurring, dry FIP treated with PI. The survival times of cats treated with PI were measured from the start of treatment and compared to survival times reported in the literature for cats with dry FIP. An improvement in survival was identified, with four long-term survivors living for over 300 days. Additionally, treating veterinarians reported a subjective improvement in the wellbeing of many of the cats (Legendre *et al.*, 2017). Though these results sound promising, they must be interpreted cautiously since neither study included a control group and definitive diagnosis of FIP with histopathology and immunostaining was not a requirement for inclusion in the studies.

A placebo-controlled study of six experimentally infected cats showed development of FIP in all cats in the control group compared to one of three undergoing treatment with anti-feline TNF-α monocloncal antibody (Doki *et al.*, 2016). The same group have demonstrated the efficacy of FCoV-derived peptides in preventing FCoV infection *in vitro*, based on a current treatment for human immunodeficiency virus in which the peptides are hypothesised to inhibit infection through virus and receptor binding (Doki *et al.*, 2015). Other compounds that were efficacious in preventing and/or treating FCoV infection *in vitro* include *Inonotus obliquus* polysaccharide (Tian *et al.*, 2016), itraconazole (Takano *et al.*, 2019a), diphyllin (Hu *et al.*, 2017), ciclosporin A (Tanaka *et al.*, 2013) and micro RNAs targeted against FCoV (Anis *et al.*, 2016). However, none of these treatments have been tested *in vivo*.

Kim et al. (2016) demonstrated the safety of the coronavirus 3C-like protease inhibitor 'GC376' in four healthy cats and showed that its presence in cell culture did not induce FIPV strain 79-1146 to evolve to overcome the inhibitory effects of the drug at 20 passages. The group went on to test the efficacy of GC376 in eight cats experimentally infected with a field strain (m3c-2) of Type 1 FCoV. Treatment was given at two different stages of disease representing early and late FIP, but in both stages the authors deemed the FIP terminal without intervention. Clinical signs and laboratory parameters consistent with FIP were reversed in six out of eight cats, and they remained healthy for the eight-month monitoring period. Interestingly, the two cats that were euthanised had severe clinical signs that, on *post mortem* examination, were attributed to severe pancreatitis and not FIP lesions. The authors were unable to explain why the cats had pancreatitis. A further study in 20 cats with naturally occurring FIP demonstrated long-term survival in seven cats treated with GC376. Six of

the seven survivors had developed acute, wet FIP and were less than 5 months old at presentation. The seventh survivor was an adult cat with only mediastinal lymph node involvement. Of the cats who succumbed to FIP despite GC376 treatment, 11 initially presented with dry FIP and all but three were older than 5 months. These results suggest that GC376 may be an effective treatment in very young cats with acute, wet FIP. Despite none of the cats in the study initially presenting with neurological signs, eight of the 13 cats who succumbed to FIP had marked neurological disease at time of death. This suggests that the drug may be selecting for neurotropic viral mutations (Pedersen *et al.*, 2017).

The nucleoside analogue 'GS-441524' was demonstrated to be effective at suppressing FIPV in cell culture, so, following a pharmacokinetic study, was tested as a treatment for FIP in 10 experimentally infected cats. Treatment reversed the signs of FIP in all 10 cats, and they remained healthy for the eight-month study follow up period (Murphy *et al.*, 2018). Thirty-one naturally infected cats were then treated for FIP with GS-441524, and, of these, 25 cats recovered completely and remained healthy for the duration of the study (at least nine months follow up) or, in the case of one cat, died of an unrelated cause (Pedersen *et al.*, 2019). GS-441524, possibly used in conjunction with GC376 or other antiviral drugs, presents an exciting novel therapeutic option for FIP.

1.8 Prevention of FIP

Prevention of FIP can be approached in two ways: prevention of FCoV transmission or prevention of the development of FIP from FCoV infection. The former is more applicable to multi-cat environments, while the latter is relevant to any FCoV-infected cat.

1.8.1 Prevention of FCoV transmission

FCoV, being such a common and contagious virus, is endemic in almost all breeding facilities, boarding catteries and shelters (Hartmann, 2005). The likelihood of FIP being present in a cattery was shown to increase with the size of the cattery (Kass, 1995), but it is unclear whether this is independent of there simply being a higher prevalence of FCoV. Though not appropriate for diagnosis of FIP, serum antibody titre and faecal RT-PCR are useful in multi-cat environments to check FCoV status, identify persistently infected individuals serving as a source of infection for other cats and monitor the impact of any changes implemented. Specific strategies to eradicate FCoV from such environments include quarantining cats that are shedding high levels of FCoV in faeces and early weaning and isolation of kittens (Addie *et al.*, 2004, Addie *et al.*, 2009, Drechsler *et al.*, 2011), though these strategies can be impractical to implement.

Addie *et al.* (2019) looked at the role of cat litter in preventing the transmission of FCoV. The authors recommend the use of a non-tracking, Fuller's earth-based litter based on its ability to prevent the virus from infecting cells *in vitro*, and possibly reduce viral transmission when used in a breeding household. As FCoV can survive for up to seven weeks in the environment and is spread *via* the faecal oral route, proper disinfection of litter boxes and all potentially contaminated surfaces is recommended to reduce transmission between individuals (Drechsler *et al.*, 2011).

1.8.2 Vaccination against FCoV

Historically the aim of vaccination has not been to prevent the cat from becoming infected when challenged with FCoV, but rather to induce a state of premonition immunity that allows the cat to contain FCoV and prevent FIP from developing (Pedersen, 2009).

Only one FCoV vaccine is in production, and it is commercially available only in the USA and some European countries: an intranasally-administered, temperature-sensitive FCoV mutant developed by passaging Type 2 FIPV strain DF2 in cell culture at 31 °C followed by ultraviolet irradiation (Christianson *et al.*, 1989). Its efficacy is controversial, with experimental and field studies reporting protection in 0 to 78% of cats (Olsen, 1993). The vaccine does not protect cats less than 16 weeks of age and is thought to be ineffective in individuals that have previously been exposed to FCoV (Addie *et al.*, 2009, Drechsler *et al.*, 2011), restricting the number of cats suitable for vaccination.

Two studies have vaccinated cats with a recombinant vaccinia virus expressing the Type 2 FCoV S protein, but in both cases ADE resulted in decreased survival rate and time in vaccinated compared to unvaccinated cats when challenged with FIPV (Vennema *et al.*, 1990, Klepfer *et al.*, 1995). Similarly, Balint *et al.* (2014) found that a recombinant FCoV vaccine based on the Type 2 FIPV strain DF2 strain, generated *via* reverse genetics, induced ADE of disease in British Shorthair cats. Interestingly, the vaccine provided complete protection against an FIPV challenge to specific pathogen free cats, illustrating the importance of host factors in determining disease course. A DNA vaccine containing the FCoV M and N genes did not confer protection against an FIPV challenge and, when administered with a cytokine-encoding plasmid, again resulted in ADE of disease (Glansbeek *et al.*, 2002).

Through reverse genetics, Haijema *et al.* (2004) generated FIPVs, based on Type 2 strain 79-1146, in which ORF3a-c, ORF7a-b or both accessory gene clusters were deleted. These recombinant FIPVs replicated well in cell culture and all displayed an attenuated phenotype *in vivo*, inducing a humoral immune response but not producing clinical signs of disease. The study found that FIPV without ORF3a-c and FIPV without OFR7a-b provided protection against an FIPV challenge in 90% of cats

without any sign of ADE, whereas FIPV where both accessory gene clusters were deleted provided no protection and in fact appeared to speed up the progression of disease compared to the control group. The authors speculated that this could reflect the latter construct's inability to induce cellular immunity. The performance of the vaccine in a non-experimental situation remains to be tested.

1.8.3 Reducing the risk of FCoV developing into FIP

In order to reduce the risk of FIP developing in a FCoV-infected individual, Addie *et al.* (2009) recommend minimising the number of cats in a household, keeping well-adapted sub-groups of no more than three cats, maintaining hygiene standards, allowing outdoor access and providing multiple litter trays that are cleaned regularly and kept away from food bowls. These measures aim to reduce stress, viral load and co-infections, all of which are known to contribute to FIP pathogenesis (Drechsler *et al.*, 2011). Since a specific stressful event such as adoption or surgery is known to precipitate FIP in many cases (Riemer *et al.*, 2015), it is logical to avoid such events in FCoV-infected cats wherever practically possible.

1.8.4 Breeding FIP-resistant cats

Hartmann (2005) recommends cessation of breeding from cats that produce two or more litters in which kittens develop FIP, based on the notion that susceptibility to FIP has a genetic basis. Recent studies have identified candidate genes that could contribute towards this susceptibility (Golovko *et al.*, 2013, Pedersen *et al.*, 2014, Wang *et al.*, 2014), but the importance of individual genes in the complex pathogenesis of FIP has been called into question (Pedersen *et al.*, 2016). On retrospective analysis of 145 cats seen at a referral hospital, Hsieh and Chueh (2014) identified a single nucleotide polymorphism (SNP) in the feline interferon-gamma (IFN-γ) gene where 20% of cats with the heterozygous genotype had FIP compared to 48% of homozygotes. This led to the development of a genetic test, now offered commercially, for 'FIP resistance'. Unsurprisingly, doubt has been cast on the validity of this test because of its apparent attribution of FIP resistance to a single gene that is not even strongly associated with FIP. Kedward-Dixon *et al.* (2019) found no association between genotype and FIP status for SNPs in the IFN-γ, TNF-α and DC-SIGN genes.

Attempts to selectively breed FIP-resistant cats found that survival following an FIPV challenge was actually decreased by positive selection (Pedersen *et al.*, 2008, Pedersen *et al.*, 2016). The authors hypothesise that this effect resulted from a loss of genetic diversity, which is in agreement with the finding that pedigree cats, who may exhibit decreased genetic diversity compared to non-pedigree cats, are more susceptible to FIP (Norris *et al.*, 2005, Pesteanu-Somogyi *et al.*, 2006, Rohrbach *et al.*, 2001).

1.9 *In vitro* propagation of FCoV

The FCoV strain Black, first isolated and identified as a Type 1 FCoV in the 1970s, was successfully grown in cell culture following animal passage and has since been maintained *in vitro* in the *Felis catus* whole fetus (FCWF) cell line (Tekes *et al.*, 2008). Regan *et al.* (2010) also describe infecting *ex vivo* feline dendritic cells, a feline-derived lung cell line named 'AKD' and Crandell feline kidney (CrFK) cells engineered to overexpress DC-SIGN with the Black strain. However, the strain is so adapted to cell culture that it can no longer be said to represent Type 1 FCoV found in the field (Tekes *et al.*, 2012, Pedersen, 2009). A novel approach from Desmarets *et al.* (2013) involved developing an immortalised feline intestinal epithelial cell line to propagate Type 1 FCoV strains 'UCD' and 'UG-FH8', which were originally obtained from clinical material and have since been passaged through live cats. The group were able to grow these viruses in cell culture, but titres obtained from this method were very low (around 10^2-10^3 PFU/ml).

The majority of *in vitro* work has been carried out on Type 2 FCoV as it can easily be propagated in cell lines expressing fAPN on their surface (Pedersen, 2014b), such as CrFK cells (Miguel *et al.*, 2002, Dewerchin *et al.*, 2005, Harun *et al.*, 2013). Conversely, a robust system for propagating field strains of Type 1 FCoV has not yet been developed (Pedersen, 2014b), likely due to the absence of cell lines that this virus can infect. This is compounded by a lack of knowledge about a Type 1 FCoV CER; this information would help to guide the search for a suitable cell line or would enable the development of a cell line stably expressing a Type 1 FCoV CER. By creating a chimeric virus comprising a Type 1 FCoV backbone with a Type 2 FCoV spike, Tekes *et al.* (2010) demonstrated that the differences observed between the serotypes *in vitro* in receptor usage, cell tropism and replication kinetics is solely down to the S protein. This suggests that a cell line such as CrFK engineered to express a Type 1 FCoV CER would likely support growth of this virus.

1.9.1 Identification of coronavirus receptors

The first coronavirus CER to be identified was that of mouse hepatitis virus (MHV). Williams *et al.* (1991) demonstrated that MHV bound to a glycoprotein in liver and intestinal brush border preparations, and that the sequence of the glycoprotein had a high homology with the human carcinoembryonic antigen family of proteins. Based on this evidence, Dveksler *et al.* (1991) reverse transcribed and PCR amplified the gene encoding murine carcinoembryonic antigen-related adhesion molecule 1 (mCEACAM1) from mouse RNA and expressed it in MHV-resistant cells, rendering them susceptible to infection with MHV and demonstrating that mCEACAM1 is a functional CER for MHV.

Porcine APN (pAPN) was first identified as a CER for TGEV when monoclonal antibodies were raised against cells susceptible to infection with TGEV, and an antibody that was able to block infection was taken forward. The protein that was precipitated from an intestinal brush border preparation by this monoclonal antibody was sequenced and identified as pAPN. Finally, pAPN was expressed in TGEV-resistant cells and shown to render them permissive to infection with TGEV (Delmas *et al.*, 1992). Similarly, human APN (hAPN) was identified as a CER for human coronavirus 229E (HCoV-229E) by demonstrating that previously uninfectable cells became infectable with HCoV-229E when engineered to express hAPN (Yeager *et al.*, 1992).

Based on the identification of APN as a CER for other Alphacoronaviruses, the role of fAPN in FCoV infection was explored. Stable expression of fAPN in FCoV-resistant mouse and hamster cell lines rendered them permissive to infection with Type 2 FCoV, demonstrating that fAPN is a functional receptor for this virus. Interestingly, fAPN was also shown to serve as a receptor for TGEV, HCoV-229E and CCoV (Tresnan *et al.*, 1996). Benbacer *et al.* (1997) then ascertained that APN is a CER for Type 2 but not Type 1 FCoV by expressing APN in FCoV-resistant cells and rendering them infectable by Type 2 FCoV only. Hohdatsu *et al.* (1998) confirmed this result by demonstrating that a monoclonal antibody raised against FCWF cells precipitated fAPN from an intestinal brush border preparation and blocked infection of FCWF by Type 2 but not Type 1 FCoV. However, FCWF cells do not normally support propagation of Type 1 FCoV so the virus used in this study must not represent a field strain. To definitively show that fAPN is not a CER for Type 1 FCoV, Dye *et al.* (2007) produced viral pseudotypes bearing the Type 1 FCoV spike protein and demonstrated that these constructs failed to recognise fAPN on three cell lines.

Most recently, a CER for MERS-CoV was identified by Raj *et al.* (2013) using the 'bait protein' method. This method utilised a chimeric protein comprising the S1 domain of the MERS-CoV spike protein and the Fc region of human IgG. This chimeric bait protein was incubated with a lysate of cells susceptible to infection, and the protein it precipitated was identified as dipeptidyl peptidase 4 (DPP4) using mass spectrometry. To demonstrate the functional relevance of DPP4 as a CER for MERS-CoV, experiments were then carried out to show that anti-DPP4 antibodies block infection of susceptible cells, and expression of DPP4 in previously MERS-CoV-resistant cells renders them permissive to infection. Since the start of this project, Zhang *et al.* (2017) have utilised a similar method to identify an interaction in chicken lung, kidney and proventriculus tissues between the avian infectious bronchitis virus (IBV) S1 protein and heat shock protein 70 (Hsp70), suggesting that the latter functions as part of the receptor complex for the virus.

1.9.2 Coronavirus reverse genetic systems

Reverse genetics is the name given to an approach whereby the function of a genetic mutation is elucidated by engineering that mutation into a genome, *in vitro*, and observing the effect of the mutation on the organism's phenotype. This is distinct from classical genetics, whereby a mutant phenotype is observed before the mutation underlying it is uncovered (Griffiths, 2000). Reverse genetic systems not only permit the study of targeted genetic changes in viral genomes, but also serve as powerful platforms for vaccine development through directed attenuation and targeted modifications (Stobart and Moore, 2014).

The first coronavirus reverse genetic systems relied on targeted genetic recombination, where the tendency of coronaviruses to swap genetic material was exploited by introducing a synthetic donor RNA and recipient virus into the same cell and allowing them to naturally recombine (Cavanagh, 2008). More recently, coronavirus reverse genetic systems have utilised the generation of full-length recombinant viruses, also known as 'infectious clones'. For a time it was unclear whether it would be technically feasible to build infectious clones of such large viruses, but three methods have shown their construction to be possible: *in vitro* ligation, bacterial artificial chromosome (BAC) systems and vaccinia virus vectors (Almazan *et al.*, 2014).

For *in vitro* ligation, full-length cDNA is assembled from a panel of contiguous fragments spanning the viral genome. The fragments are generated either by RT-PCR from virus or synthesised *de novo*. Sequences are introduced to incorporate a T7 RNA polymerase promotor at the 5′ end and a poly(A) tail at the 3′ end of the construct. Older methods used silent mutations to insert restriction sites and avoid T7 transcription termination signals, but this has been overcome using Type IIS restriction enzymes. Type IIS enzymes recognise asymmetrical sites and cleave 1-4 nucleotides away from the recognition sequences, leaving overhangs that can only anneal with complementary DNA generated at an identical site. If fragments are designed so that the recognition sequence is cleaved from the fragment, the exact viral sequence is left behind resulting in seamless, directional cloning. The main advantage of *in vitro* ligation is that it is relatively simple and fast, and that cDNA fragments incorporating mutations of interest can easily be interchanged. Additionally, constructs generated through in vitro ligation can later be cloned into BAC and vaccinia systems if the correct flanking sequences are present (Almazan *et al.*, 2014). This method, originally pioneered by Yount *et al.* (2000) to create a TGEV infectious clone, has been used to produce many recombinant coronaviruses but not FCoV.

In a BAC system, viral genome fragments are assembled into a very low copy number synthetic plasmid which is maintained in bacterial cells. The advantages of this system are stable maintenance

of large DNA fragments and unlimited production of the cDNA clone. Additionally, features such as a CMV promoter or a synthetic poly(A) tail can be engineered into the plasmid, allowing recovery of infectious virus without an *in vitro* transcription step (Almazan *et al.*, 2014). This system has been successfully used to build infectious clones of coronaviruses including a Type 2 FIPV (Balint *et al.*, 2012).

In a vaccinia virus system, coronavirus cDNA fragments are assembled *in vitro* then ligated into vaccinia virus DNA. The ligation reaction is transfected into mammalian cells and recombinant vaccinia viruses containing the coronavirus cDNA are rescued. Purified vaccinia virus DNA is cleaved so that the coronavirus cDNA insert drops out, and the insert is transcribed *in vitro*. The advantages of this system are similar to those of the BAC system: the cDNA clone is maintained stably and can be produced infinitely through infection of cells with the recombinant vaccinia virus. Additionally, vaccinia virus-mediated homologous recombination can be used to easily introduce modifications into the coronavirus genome. However, generation of recombinant vaccinia viruses is time-consuming because it involves all the work of *in vitro* ligation with additional steps to ligate the construct into vaccinia virus then remove it again for transcription (Almazan *et al.*, 2014). Since the start of this project, the vaccinia virus method has been utilised by Ehmann *et al.* (2018) to build the first Type 1 FCoV infectious clone representing a field strain of the virus.

Development of a cell line supporting the *in vitro* propagation of field strains of Type 1 FCoV would be an important component of a Type 1 FCoV reverse genetic system, as it would enable *in vitro* recovery of recombinant viruses.

1.10 Aims of the project

Despite recent progress in the field of treatment, most cases of FIP are fatal and cause a great deal of suffering, not only to the affected cat but to those who care for that cat. A better understanding of how FIP develops from FCoV, and particularly biologically relevant field strains of Type 1 FCoV, will be indispensable for the rational design of vaccines, treatments and prevention strategies.

That is why **the overall aim of this project is to establish a reverse genetic system based on a field strain of Type 1 FCoV**. Such a system would allow the effect of a given mutation on the phenotype of the virus to be elucidated, providing insight into FIP pathogenesis. In order to achieve that overall aim, the specific aims of this project are:

1. To develop cDNA infectious clones corresponding to full-length Type 1 FCoV, chimeric Type 1 FCoV with a Type 2 S gene and sub-genomic replicons, and to test the ability of the clones to replicate in cell culture.
2. To identify a CER for Type 1 FCoV, for the purpose of developing a cell line that can support *in vitro* propagation of the virus,
3. To explore the use of the feline IFN-γ ELISpot assay in measuring T cell response to viral epitopes, for the purpose of identifying mutations that may be of use for engineering attenuated recombinant viruses.

In this project, cDNA infectious clones will form the basis of the reverse genetic system. A specific aim of this project is therefore **to develop cDNA infectious clones corresponding to full-length Type 1 FCoV, chimeric Type 1 FCoV with a Type 2 S gene and sub-genomic replicons, and to test the ability of the clones to replicate in cell culture**. At the commencement of this project there was no report of a reverse genetic system based on a field strain of Type 1 FCoV in the literature, so its development would have been novel. Ehmann *et al.* (2018) have since published a report detailing their establishment of such a system, and this will be discussed further in chapters 3 and 6.

A major limitation with the study by Ehmann *et al.* (2018), and something that has long impeded FCoV research, is absence of a cell line in which Type 1 FCoV can be propagated. The authors had no choice but to use their recombinant Type 1 FCoV to directly infect live cats, when an *in vitro* characterisation stage may have been more informative. Desmarets *et al.* (2013) developed an immortalised feline intestinal epithelial cell line in which Type 1 FCoV can be grown, but this cell line is unavailable to other researchers. A way to rescue recombinant virus *in vitro* would certainly be important to this project, since we plan to use the reverse genetic system to explore the effect of viral mutations on viral phenotype *in vitro*.

A specific aim of this project is therefore **to identify a CER for Type 1 FCoV**. Knowledge of a CER that Type 1 FCoV uses would allow this receptor to be stably expressed in a feline cell line. Such a cell line should support propagation of this serotype, and therefore enable rescue of recombinant Type 1 FCoV and its characterisation *in vitro*.

An important potential use of the reverse genetic platform is the rational design and production of a vaccine against FIP. Reverse genetics has been used to generate candidate vaccines against FIP but these have been based on Type 2 FIPVs (Haijema *et al.*, 2004, Balint *et al.*, 2014) so their biological relevance is questionable. An effective vaccine against FIP would likely have to mimic natural FIP resistance, where a cat becomes infected with a relatively avirulent strain of Type 1 FCoV then is able to contain that infection and prevent it from developing into FIP. The factors that contribute to this disease outcome are not entirely understood, but a robust cellular immune response is thought to play an important part (Pedersen, 2014b).

Based on this idea, another specific aim of this project is **to explore the use of the feline IFN-γ ELISpot assay in measuring T cell response to viral epitopes**. Within this project, a feline ELISpot assay that measures the production of IFN-γ will be optimised and used to analyse the effect of a specific FCoV mutation in a known T cell epitope. T cells will be stimulated by exposing peripheral blood mononuclear cells (PBMC) to peptides with and without the mutation. IFN-γ ELISpot represents a tool for identifying the FCoV epitopes that are best able to elicit a cellular immune response, and with this knowledge any candidate vaccines generated by the reverse genetic platform could be designed to include these epitopes.

2 Materials and methods

2.1 Materials

2.1.1 Cell culture media

Cell line medium	Dulbecco's modified Eagle medium (DMEM) with high glucose, GlutaMAX supplement and pyruvate (Thermo-Fisher, Massachusetts, USA), supplemented with 1% (v/v) penicillin-streptomycin (10,000 U/ml; Thermo-Fisher), 100 µM non-essential amino acids (Sigma-Aldrich, Gillingham, UK) and 10% (v/v) heat-inactivated fetal bovine serum (FBS; Thermo-Fisher)
Epithelial cell medium	DMEM/Nutrient F-12 Ham, supplemented with 2 mM L-glutamine (Thermo-Fisher), 1% (v/v) penicillin-streptomycin, one HCM SingleQuots growth supplement pack (Lonza, Basel, Switzerland) and 5% (v/v) heat-inactivated FBS
PBMC medium	RPMI 1640 (Lonza), supplemented with 2 mM L-glutamine, 1% (v/v) penicillin-streptomycin and 10% (v/v) heat-inactivated FBS
Primary culture medium	Advanced DMEM/F12 (Thermo-Fisher) supplemented with 2 mM L-glutamine, 1% (v/v) penicillin-streptomycin and 20% (v/v) heat-inactivated FBS

2.1.2 Bacterial culture media

Super optimal broth with catabolite repression (SOC) medium	20.0 g/l tryptone, 5.0 g/l yeast extract, 0.5 g/l NaCl, 5.0 g/l $MgSO_4$, 0.2 g/l KCl, 2.03 g/l $MgCl_2$ and 3.6 g/l glucose
Luria broth (LB) agar plates	10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl and 15 g/l agar
LB medium	10 g/l tryptone, 5 g/l yeast extract and 10 g/l NaCl

2.1.3 Solutions, buffers and detergents

Phosphate buffered saline (PBS)	138 mM NaCl, 2.7 mM KCl, 8 mM Na2HPO4, 1.5 mM KH2PO4
Tris-borate-EDTA (TBE)	89 mM Tris base, 89 mM boric acid, 2 mM EDTA
Tris-acetate-EDTA (TAE)	40 mM Tris base, 20 mM acetic acid, 1 mM EDTA
Annealing buffer	50mM Tris-HCl pH 8.0, 100 mM NaCl, 1mM EDTA
SDS-PAGE sample buffer; 2x	7 ml 0.5 M Tris-HCl pH6.8, 3 ml Glycerol, 1.2 g SDS, 1.2 mg Bromophenol blue, made up to 30 ml with deionised water
Coomassie Blue staining solution	0.575 g Coomassie Brilliant Blue R-250, 200 ml methanol, 50 ml glacial acetic acid, made up to 250 ml with deionised water
Destaining solution	100 ml methanol, 50 ml glacial acetic acid, 350 ml deionised water
Transfer buffer	100 ml methanol, 12.5 g Tris base, 5.63 g glycine, made up to 500 ml with deionised water
Cell surface lysis buffer	5 ml PBS containing 1 cOmplete Mini EDTA-free Protease Inhibitor Cocktail tablet (Roche, Basel, Switzerland) and 0.3% (w/v) n-Dodecyl ß-D-maltopyranoside (Thermo-Fisher)
Whole cell lysis buffer	20 mM Tris HCl pH 8, 137 mM NaCl, 1% (v/v) Triton X-100, 2 mM EDTA with 1 cOmplete Mini EDTA-free Protease Inhibitor Cocktail per 5 ml
Elution buffer	70% (v/v) water, 5% (v/v) β-mercaptoethanol, 25% (v/v) 4x SDS-PAGE sample buffer

2.1.4 Antibodies

Antibody	**Manufacturer and catalogue number**	**Application**	**Concentration**
Rabbit anti-human IgG Fc antibody	Thermo-Fisher; 31142	Primary antibody in western blot	1 in 1000
Goat anti-rabbit HRP-conjugated antibody	Vector Laboratories, California, USA; PI-1000	Secondary antibody in western blot	1 in 1000
Goat anti-rabbit HRP-conjugated antibody	Agilent, California, USA; P0448	Secondary antibody in western blot	1 in 5000
Rabbit anti-HSPA1A antibody	Abcam, Cambridge, UK; ab182844	Primary antibody in western blot Primary antibody in immunofluorescence	1 in 2000 1 in 100
Rabbit anti-human IgG CF 633-conjugated antibody	Sigma-Aldrich; SAB4600149	Secondary antibody in flow cytometry Secondary antibody in immunofluorescence	1 in 100 to 1 in 1000 1 in 1000
Mouse anti-FLAG	Sigma-Aldrich; F1804	Primary antibody in western blot Primary antibody in immunofluorescence	1 in 1000 1 in 500
Goat anti-human IgG DyLight 488-conjugated antibody	Abcam; ab97003	Secondary antibody in immunofluorescence	1 in 250 to 1 in 500
Mouse anti-coronavirus (FIPV3-70) antibody	Santa Cruz, Dallas USA; sc-65653	Primary antibody in immunofluorescence	1 in 100
Goat anti-mouse Alexa Fluor 488-conjugated antibody	Thermo-Fisher; A11029	Secondary antibody in immunofluorescence	1 in 700

Goat anti-mouse Alexa Fluor 568-conjugated antibody	Thermo-Fisher; A11031	Secondary antibody in immunofluorescence	1 in 1000
Goat anti-rabbit Alexa Fluor 568-conjugated antibody	Thermo-Fisher; A11036	Secondary antibody in immunofluorescence	1 in 1000

2.2 Cell culture methods

2.2.1 Cell lines

Cell lines used were human embryonic kidney 293T (HEK293T; CRL-3216, American Type Culture Collection, Virginia, USA), Crandell feline kidney (CrFK; CCL-94, American Type Culture Collection), DH82 (CRL-10389, American Type Culture Collection), feline embryo-A (FE-A) (Jarrett *et al.*, 1973), baby hamster kidney (BHK-21; CCL-10, American Type Culture Collection) and L-WRN (CRL-3276, American Type Culture Collection) cells. Cells were were grown at 37 °C in a humidified incubator with 5% CO_2 on cell culture plates or flasks. Cells were passaged when approximately 90% confluent by washing once with phosphate buffered saline without calcium and magnesium (PBS; Thermo-Fisher), detaching from the flask with 0.05% (v/v) trypsin-EDTA in PBS (Thermo-Fisher), adding cell line medium and harvesting the cells by centrifugation at 300 x *g* for 3 minutes. The cell pellet was resuspended in cell line medium and an aliquot was used to seed a new culture vessel.

2.2.2 Intestinal epithelial cells

Gut samples were obtained from cats euthanased for clinical reasons not related to this study. Ethical approval was awarded and owner consent obtained for post-mortem examination and collection of tissue samples for research (University of Bristol veterinary investigation number VIN/14/013) and the use of these samples in cell culture (University of Bristol veterinary investigation number VIN/16/033). As soon as possible after euthanasia, approximately 1 cm lengths of jejunum, ileum and colon were removed and placed in cold, neat histidine tryptophan ketoglutarate (Pharmapal Ltd, Elstree, UK) on ice. Visible fat and connective tissue was removed from the intestine, it was opened longitudinally, then it was washed three times in PBS containing 1% (v/v) penicillin-streptomycin (10,000 U/ml; Thermo-Fisher) and 1 µg/ml amphotericin-B (Sigma-Aldrich).

The epithelial cells were isolated from intestinal tissue as described by Desmarets *et al.* (2013) with some modifications, whereby the tissue was digested with 0.4 mg/ml collagenase (Thermo-Fisher) and 1.2 mg/ml dispase (Sigma-Aldrich) for 1 hour with a refreshment of medium in the middle of the incubation time. The digested intestinal mucosa was scraped with a sterile scalpel blade into warm DMEM (Thermo-Fisher) supplemented with 1.2 mg/ml dispase for 10 minutes with constant pipetting. The resulting cell suspension was passed through a 70 µm cell strainer and centrifuged at 500 x *g* for 10 minutes. The cell pellet was resuspended in epithelial cell medium and grown at 37 °C in a humidified incubator with 5% CO_2 on cell culture plates or flasks.

2.2.3 Peripheral blood mononuclear and monocyte-derived cells

Feline whole blood, from which PBMC were isolated, was obtained as surplus from routine venepuncture of cats undergoing diagnostic testing for an unrelated condition, or *post mortem* from cats euthanased for clinical reasons not related to this study. Ethical approval was awarded and owner consent obtained for collection of surplus material from patients (University of Bristol veterinary investigation number VIN/14/020), for *post mortem* examination and collection of tissue samples for research (University of Bristol veterinary investigation number VIN/14/013) and the use of these samples in cell culture (University of Bristol veterinary investigation number VIN/16/033). Approximately 1 ml of blood was collected into tubes containing EDTA (whichever used normally by the veterinary practice in which the blood was being taken), stored at room temperature and processed as soon as practically possible after collection (usually within two days).

After removing the plasma, blood was resuspended in 6 ml PBMC medium. This was carefully layered onto 6 ml of Lymphoprep (Axis-Shield PoC AS, Oslo, Norway) in a second 15 ml Falcon tube, using a Pasteur pipette, before centrifugation at 1,000 x *g* for 20 minutes with the brake off. A Pasteur pipette was then used to aspirate the buffy coat and transfer it to a third 15 ml Falcon tube. The buffy coat was topped up to 14 ml with PBMC medium before gentle inversion and centrifugation at 700 x *g* for 10 minutes. The supernatant was discarded, and the pellet resuspended in 5 ml PBMC medium.

The resulting PBMC were either used in an ELISpot assay (section 2.4.1), applied to a glass slide by cytocentrifugation at 1,000 rpm for 5 minutes using a Cytospin 4 (Thermo-Fisher) for use in an immunofluorescence assay (section 2.4.10) or cultured to obtain monocyte-derived macrophages. For the latter, 5-6 x 10^5 live (as determined by mixing an equal volume of cell suspension and trypan blue, transferring into a counting chamber and counting only cells that did not take up the stain) PBMC in 1 ml PBMC medium were seeded into each well of a 24 well plate containing a glass coverslip. After 24 hours, the cells were gently washed twice with warm medium to remove non-adherent cells. The resulting cell population was cultured in PBMC medium for a further 4 days to obtain monocyte-derived macrophages.

2.2.4 Intestinal organoids

2.2.4.1 Making conditioned medium

Conditioned medium was made by following the 'small scale preparation' protocol of Miyoshi and Stappenbeck (2013). L-WRN cells were grown in cell line medium at 37 °C in a humidified incubator with 5% CO_2 in cell culture flasks, in the presence of 500 µg/ml G418 (Thermo-Fisher) and 500 µg/ml

hygromycin B (Thermo-Fisher). Cells were passaged when confluent by washing once with PBS, detaching from the flask with 0.05% (v/v) trypsin-EDTA in PBS, adding cell line medium and using aliquots of the cell suspension to seed new flasks. Cells were grown in cell line medium at 37 °C in a humidified incubator with 5% CO_2 (without G418 and hygromycin) until overconfluent, then the cell line medium was replaced with primary culture medium. The medium was collected and replaced every 24 hours, centrifuged at 2000 x *g* for 5 minutes to remove cells, pooled and diluted with an equal volume of primary culture medium. The resulting conditioned medium was stored in aliquots at -20 °C and kept for two weeks once thawed.

2.2.4.2 Isolating feline intestinal crypts

Gut samples were obtained from cats euthanased for clinical reasons not related to this study. Ethical approval was awarded and owner consent obtained for post-mortem examination and collection of tissue samples for research (University of Bristol veterinary investigation number VIN/14/013) and the use of these samples in cell culture (University of Bristol veterinary investigation number VIN/16/033).

As soon as possible after euthanasia, an approximately 1 cm length of ileum was removed and placed in PBS on ice. The tissue was trimmed to remove connective tissue and fat, then opened longitudinally. The lumen was washed with cold PBS, then the whole piece of tissue was shaken vigorously in a tube containing cold PBS. The tissue was cut into approximately 1 cm^3 pieces and washed again with cold PBS. The tissue was placed in cell line medium, then the luminal surface was gently scraped with a sterile scalpel blade to remove villi. The tissue pieces were minced using scissors, then incubated in 2 mg/ml collagenase (Thermo-Fisher) in cell line medium for 90 minutes with frequent agitation. The resulting digest was passed through a 70 µm cell strainer (Thermo-Fisher) and the filtrate centrifuged at 50 x *g* for 5 minutes. The pelleted cells (including crypt cells) were washed twice in cell line medium, resuspended in Matrigel matrix (growth factor reduced, phenol red free; catalogue number 356234; Corning, New York, USA) and pipetted on to a cell culture plate. The Matrigel domes were set at 37 °C for at least 5 minutes then submerged in conditioned medium supplemented with 10 µM Y-27632 (Enzo Life Sciences, Exeter, UK), 10 µM SB-43154 (Stemcell, Vancouver, Canada) and 10 mM nicotinamide (Sigma-Aldrich).

2.2.4.3 Growing and passaging intestinal organoids

Murine intestinal organoids (a gift from Dr Rhiannon Jenkinson, University of Bristol) were cultured at 37 °C in a humidified incubator with 5% CO_2 in 48-well plates and passaged every 7 days. They were liberated from Matrigel matrix by adding 1 ml ice cold PBS, then mechanically dissociated by

pipetting vigorously and passing through a 27 G needle. The organoids were transferred to 5 ml ice cold PBS and centrifuged at 150 x *g* for 10 minutes at 4°C.

Feline intestinal organoids (isolated in-house as described in section 2.2.4.2 or a gift from Dr Michael Behnke, University of Louisiana, USA) were cultured at 37 °C in a humidified incubator with 5% CO_2 in 24-well plates and passaged every 3-7 days. They were liberated from Matrigel matrix by adding 200 µl 0.125% (v/v) trypsin-EDTA in PBS and suspending the matrix by pipetting with a 1,000 µl pipette, then incubated at 37 °C for 5 minutes. 500 µl cell line medium were added and the organoids were transferred to 5 ml cell line medium and centrifuged at 200 x *g* for 5 minutes.

In both cases, the supernatant was taken off except for 200 µl, and the organoids were resuspended and transferred to a 1.5 ml microcentrifuge tube. Muramyl dipeptide (Bio-Techne, Abingdon, UK) was added to a final concentration of 10 µg/ml and this was incubated at room temperature for 10 minutes. 1 ml cell line medium was added, and the organoid suspension was centrifuged at 200 x *g* for 5 minutes. The supernatant was taken off and the organoids were resuspended in Matrigel matrix. 20 µl domes were pipetted on to a new plate, and these were set at 37 °C for at least 5 minutes. 350-500 µl medium were added per well. For feline organoids the medium was conditioned medium supplemented with 10 µM Y-27632, 10 µM SB-43154 and 10 mM nicotinamide. For murine organoids the medium was conditioned medium supplemented with 10 µM Y-27632 and 10 µM SB-43154, IntestiCult Organoid Growth Medium Human (Stemcell) or IntestiCult Organoid Growth Medium Mouse (Stemcell).

2.3 DNA and RNA techniques

2.3.1 Plasmid design and production

2.3.1.1 Bait protein plasmids

Quirke (2016), a previous Master of Research student in the laboratory, cloned DNA inserts encoding Type 1 and Type 2 bait proteins (~3,100 bp) into the mammalian expression vector pCAGGS (GenBank accession number LT727518.1). The bait protein DNA inserts, designed by Professor Stuart Siddell (University of Bristol), encode the S1 domain of Type 1 FCoV 80F strain (GenBank accession number KP143511.1) or Type 2 FCoV 79-1146 (GenBank accession number DQ010921.1) fused to the Fc region of human IgG. The S1 domains were codon optimized for translation in human cells, putative splice donor and receptor sites were removed and restriction enzyme cloning site sequences were added onto each end. Bait protein coding sequences are in Appendix A

Used as a positive control for transfection, pmaxGFP (Lonza) is a 3,486 bp vector containing a green fluorescent protein gene under the control of a cytomegalovirus promoter.

2.3.1.2 Infectious clone plasmids

cDNA fragments comprising the Type 1 FCoV 80F strain, Type 2 FCoV 79-1146 strain S gene, green fluorescent protein/puromycin (GFP puro) resistance coding sequence and *Renilla* luciferase coding sequence were designed using EMBOSS Sixpack and Microsoft Word software. Fragment sequences were submitted to Twist Bioscience (San Francisco, USA) for synthesis and cloning into the 2,221 bp vector pTwist 31 pMB1-Amp-UCS1.2. Fragment sequences are in Appendix A.

Some unstable or potentially toxic sequences were cloned into the low copy number vector pWSK29 (GenBank accession number AF016889.1).

Additional oligonucleotides were designed using Microsoft Word and NEBCutter V2.0 software (New England BioLabs (NEB), Massachusetts, USA), and were synthesised by Eurofins Genomics (Ebersberg, Germany). The oligonucleotide sequences are in Appendix A.

2.3.1.3 HSPA1A plasmid

A translated BLAST (NCBI, Bethesda, USA) search was carried out based on the amino acid sequence of the feline HSPA1A protein (Uniprot accession number A0A337RXE8_FELCA), which identified LOC105260573 as the encoding gene. The coding sequence was identified using ExPASy translate tool (Swiss Institute of Bioinformatics, Lausanne, Switzerland), edited to incorporate a 3xFLAG tag and restriction endonuclease sites (*Not* I and *Eco* RI) using Microsoft Word and checked for restriction endonuclease sites using NEBCutter V2.0 software.

The HSPA1A construct was cloned into the pEX-A258 vector by Eurofins Genomics, then subcloned into the mammalian expression vector pcDV4 (Hannemann *et al.*, 2013).

2.3.2 Bacterial transformation and subsequent harvesting of plasmid DNA

Alpha select chemically competent cells (Bioline, London, UK) or 5-alpha competent *Escherichia coli* (NEB) were incubated with 5 ng of plasmid DNA on ice for 30 minutes. The cells were then heat shocked at 42 °C for 30 seconds before another incubation on ice for 2 minutes. 950 µl SOC medium was added and the cells incubated at 37 °C for 1 hour with shaking (220 rpm) in an orbital shaker, then the resulting bacterial suspension was spread on LB agar plates containing 100 µg/ml ampicillin (Sigma-Aldrich) and incubated at 28-37 °C until colonies reached the desired size.

Colonies were picked and inoculated into LB medium containing 100 µg/ml ampicillin before incubation at 28-37°C with shaking (220 rpm) in an orbital shaker. Plasmid DNA was extracted from 2 ml, 50 ml and 200 ml of bacterial culture using a GeneJET Plasmid Miniprep, GeneJET Plasmid Midiprep or PureLink HiPure Plasmid Filter Maxiprep Kit (Thermo-Fisher) respectively, following the manufacturer's instructions. When dealing with low copy number plasmids, double the volume of bacterial culture was processed.

2.3.3 PCR

10 ng of DNA to be amplified by PCR were incubated with 10 µl OneTaq standard reaction buffer (5x; NEB), 200 µM deoxyribonucleotide triphosphates (dNTPs), 0.2 µM each of forward and reverse primer and 2.5 units OneTaq DNA polymerase (NEB) in a total volume of 50 µl.

When a DNA fragment >5,000 bp was to be amplified, 10 ng of DNA were incubated with 10 µl Phusion HF buffer (5x; NEB), 200 µM dNTPs, 0.5 µM each of forward and reverse primer and 1 unit Phusion DNA polymerase (NEB) in a total volume of 50 µl.

For long-range PCR, 10 ng of DNA were incubated with 25 µl GoTaq Long PCR Master Mix (2x; Promega, Madison, USA) and 0.5 µM each of forward and reverse primer in a total volume of 50 µl.

In all cases, the volume was brought up to 50 µl with deionised water, and a PCR programme with step lengths and times appropriate to the primers and enzymes was carried out in a GS1 thermocycler (G-Storm, Somerset, UK). See Appendix B for the thermocycler programmes used.

A proportion of the resulting DNA was analysed by electrophoresis through a 0.7-1.0% TBE or a 0.4% TAE agarose gel (section 2.3.6). The remaining DNA was purified using a GeneJet PCR Purification Kit (Thermo-Fisher) following the manufacturer's instructions, or by phenol chloroform extraction and ethanol precipitation (section 2.3.12).

PCR primers were chosen manually and checked using Multiple Primer Analyzer software (Thermo-Fisher). For a list of PCR primers, see Appendix A.

2.3.4 RT-PCR

RNA was reverse transcribed and amplified by PCR in a one-step RT-PCR reaction. 10-50 ng of RNA was incubated with 25 µl Platinum™ SuperFi™ RT-PCR Master Mix (2x; Thermo-Fisher), 10 µM each of forward and reverse primer and 0.5 µl SuperScript™ IV RT Mix (Thermo-Fisher). The volume was brought up to 50 µl with deionised water and, following an incubation for 10 minutes at 50 °C, a PCR programme with step lengths and times appropriate to the primers and enzymes (Appendix B) was carried out in a GS1 thermocycler.

2.3.5 Restriction enzyme digestion

Appropriate restriction enzymes (NEB) were used to digest DNA in a total volume of 15-50 µl for 1-2 hours in the buffer and at the temperature specified by the manufacturer.

The resulting products were analysed by electrophoresis through a 0.7-1.0% TBE agarose gel (section 2.3.6).

2.3.6 Agarose gel electrophoresis

Samples containing DNA were diluted with nuclease-free water and 6x gel loading dye (NEB) in an appropriate volume, and electrophoresed through a 0.7-1.0% TBE agarose gel (0.7-1.0% (w/v) UltraPure agarose (Thermo-Fisher) in TBE) at 100 V for 40 minutes, or a 0.4% TAE agarose gel (0.4% (w/v) agarose in TAE) at 15 V for 24 hours. Electrophoresis was performed in 1x TBE or TAE buffer. Both the gel and buffer contained 0.5 µg/ml ethidium bromide (Sigma-Aldrich).

Following electrophoresis, the gel was imaged on a UVP BioDoc-ITTM System Ultraviolet transilluminator (UVP, California, USA) to estimate the size of the DNA fragment(s) compared to a GeneRuler™ 1 kb Plus (Thermo-Fisher) or λ DNA-Mono Cut Mix (NEB) DNA Ladder.

2.3.7 DNA gel extraction

DNA that had undergone restriction digestion and agarose gel electrophoresis was visualised briefly using the UVP BioDoc-ITTM System Ultraviolet or Dark Reader (Clare Chemical Research, Colorado, USA) transilluminator, and the band corresponding to the desired product was excised from the gel. DNA was extracted from the gel slice using a GeneJET Gel Extraction Kit (Thermo-Fisher) following the manufacturer's instructions.

2.3.8 Antarctic phosphatase digestion

Approximately 3 µg DNA digested by restriction enzyme was incubated with 15 units Antarctic Phosphatase (NEB) and 6 µl Antarctic Phosphatase buffer (10x; NEB). The volume was brought up to 60 µl with deionised water and incubated at 37 °C for a further 30 minutes.

The product was purified using a GeneJet PCR Purification Kit (Thermo-Fisher) following the manufacturer's instructions.

2.3.9 NEBuilder HiFi DNA assembly

DNA fragments with 25 bp overlapping regions were ligated together using the NEBuilder HiFi DNA assembly protocol as described by the manufacturer (NEB). Briefly, fragments were incubated together in a 1:2 vector:insert molar ratio, in a maximum volume of 10 µl and a maximum molarity of 0.2 pmols. To this, 10 µl NEBuilder HiFi DNA Assembly Master Mix was added and the volume was brought up to 20 µl with deionised water. The mixture was incubated at 50 °C for 1 hour. 2 µl of the reaction mixture was used to transform chemically competent bacterial cells, following the procedure described in section 2.3.2.

2.3.10 Oligonucleotide ligation

Oligonucleotides were diluted to 40 ng/µl in annealing buffer and incubated as complementary pairs at 95 °C for 5 minutes. The reaction was cooled gradually to room temperature over 45 minutes to allow oligonucleotide annealing.

2.3.11 T4 DNA ligation

2.3.11.1 Ligating inserts into vectors

DNA fragments that had undergone restriction digestion or oligonucleotide ligation to form complementary oligonucleotide pairs with cohesive ends were incubated together in a 1:3 to 1:6 vector:insert molar ratio, with 2 µl T4 DNA ligase buffer (10x; NEB) and 1 µl T4 DNA ligase (NEB). The volume was brought up to 20 µl with deionised water. This was incubated at room temperature for 1 hour before enzyme inactivation at 65 °C for 10 minutes.

2.3.11.2 Ligating full-length infectious clones

DNA 'units' collectively comprising full-length infectious clones were incubated together in an equimolar ratio and approximately 5.5 µg total DNA, with 20 µl T4 DNA ligase buffer and 20 µl T4 DNA ligase. The volume was brought up to 200 µl with deionised water. This was incubated at 16 °C overnight.

A proportion of the ligation reaction was analysed by electrophoresis through a 0.4% TAE agarose gel. The remainder of the ligation reaction was purified by phenol chloroform extraction and ethanol precipitation (section 2.3.12) and used as a template for transcription.

2.3.12 Phenol chloroform extraction and ethanol precipitation

One volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added to one volume of a DNA containing solution and this was shaken by hand for 2 minutes before centrifugation at 13,000 × *g* for 5 minutes. The upper aqueous phase was removed and set aside, one volume of deionised water was added to the original sample and this was shaken by hand for 2 minutes before centrifugation at 13,000 × *g* for 5 minutes. The upper aqueous phase was again removed and set aside.

DNA was precipitated from the aqueous phase by adding a 0.1 volume of 3 M sodium acetate (pH 5.2) followed by 2.5 volumes of 100% ethanol. The solution was incubated at -20 °C for at least 30 minutes and the precipitate collected by centrifugation at 13,000 x *g* at 4 °C for 20 minutes. The pellet was washed with 70% ethanol, collected again by centrifugation and finally dried and resuspended in nuclease-free water.

2.3.13 DNA Sequencing

2.3.13.1 Sanger sequencing

DNA was submitted to Eurofins Genomics for Sanger sequencing. Sequencing data was compiled and analysed in BioEdit version 7.0.5.3 software. For a list of sequencing primers, see Appendix A.

2.3.13.2 Next generation sequencing

cDNA amplicons generated by RT-PCR were pooled in an equimolar fashion and submitted to the Genomics Facility, University of Bristol, for next-generation sequencing. Samples were prepared for sequencing using a TruSeq Nano DNA Library Preparation kit (Illumina, California, USA) and were sequenced using a MiSeq Version 2 2x250bp Nano platform (Illumina).

Reads were aligned against a reference genome using Bowtie2 software accessed through the Galaxy online bioinformatics suite. Alignments were visualised using Integrative Genomics Viewer (IGV) software.

2.3.14 RNA transcription

Transcription of RNA from DNA template was carried out using a mMessage mMachine T7 kit (Thermo-Fisher) following the protocol specified by the manufacturer with minor modifications. Briefly, 0.1-10 µg template DNA was incubated with 10 µl NTP/CAP (2x), 2 µl reaction buffer (10x), 2

µl enzyme mix and, when transcribing full-length infectious clones, 3 µl GTP (final cap analog:GTP ratio of 1:1). The volume was brought up to 20 µl with nuclease-free water. This was incubated at 37 °C for 2 hours. 1 µl of TURBO DNase was added and this was incubated at 37 °C for 15 minutes.

The resulting RNA was purified by lithium chloride precipitation as specified by the kit manufacturer and was analysed for yield and purity on a NanoDrop 1000 spectrophotometer (Thermo-Fisher).

2.3.15 Denaturing agarose gel electrophoresis

5 ml 3-(N-morpholino)propanesulfonic acid (MOPS) buffer (10x; Thermo-Fisher), 9 ml 37% formaldehyde (Sigma-Aldrich) and 25 µg ethidium bromide were added to 0.5 g agarose dissolved in 36 ml nuclease-free water. This gel was poured and set in a tray, then assembled in an electrophoresis tank submerged in 1x MOPS buffer containing 50 µg ethidium bromide per 100 ml of buffer.

RNA samples were prepared by adding one volume of NorthernMax Formaldehyde Load Dye (Thermo-Fisher) and incubating at 70 °C for 10 minutes. Ethidium bromide was added directly to the samples to a final concentration of 10 µg/ml, and the samples were run through the gel at 100 V until the bromophenol blue marker had migrated two thirds of the gel's length. The gel was destained in nuclease-free water overnight before RNA bands were visualised using a UVP BioDoc-ITTM System Ultraviolet transilluminator.

2.3.16 RNA transfection into mammalian cells

Cells were grown in cell culture plates in cell line medium at 37 °C and 5% CO_2 for ~24 hours (cells at 60-90% confluency). The RNA transfection mixture was prepared using the TransMessenger Transfection Reagent kit (Qiagen, Venlo, the Netherlands) as specified by the manufacturer, using an RNA:TransMessenger Transfection Reagent ratio of 1:4. Medium was removed from the cells, they were washed once with PBS and the transfection mixture was added dropwise. This was incubated at 37 °C and 5% CO_2 for 3 hours, then the transfection mixture was removed. The cells were washed once with PBS and an appropriate volume of cell line medium was added back into the wells.

Following transfection with the replicon construct, BHK or CrFK cells were recovered for 24 hours before being transferred to a 25 cm^2 cell culture flask. After another 24 hours, 1.5 µg/ml or 5 µg/ml puromycin (Sigma-Aldrich) was added to the medium to select for successfully transfected BHK and CrFK cells respectively.

Following transfection with the infectious clone construct, the culture supernatant was harvested after 2 days and either stored at -80 °C for analysis or passaged on to CrFK cells (passage 1). The

cells, which had been grown on glass coverslips, underwent an immunofluorescence assay (section 2.4.10). After 5 days, the passage 1 culture supernatant was harvested and either stored at -80 °C for analysis or passaged on to CrFK cells (passage 2). After 5 days, the passage 2 culture supernatant was harvested and stored at -80 °C for analysis.

2.3.17 DNA transfection into mammalian cells

For bait protein production, HEK293T cells were seeded in 6-well plates at 3 x 10^5 cells in 2.5 ml cell line medium per well. For pHSPA1A transfection, CrFK cells were seeded in either 6-well plates at 1 x 10^5 cells in 2.5 ml cell line medium per well, or 24-well plates containing glass coverslips at 2 x 10^4 cells in 1 ml cell line medium per well.

After culturing at 37 °C and 5% CO_2 for ~24 hours (cells at 60-90% confluency), the cell line medium was removed and the transfection mixture (10 µl Lipofectamine 2000 (Thermo-Fisher) and 4 µg sterile plasmid DNA per transfection reaction for 6-well plates or 2 µl Lipofectamine 2000 and 0.5 µg sterile plasmid DNA per transfection reaction for 24-well plates, prepared in Opti-MEM Reduced Serum Medium (Thermo-Fisher) according to the manufacturer's instructions) was added dropwise. This was incubated at 37 °C and 5% CO_2 for 3 hours. The transfection mixture was then carefully removed and 2.5 ml Pro293A-CDM (Lonza) was added per well for bait protein production, and 2.5 ml (6-well plate) or 1 ml (24-well plate) cell line medium per well for pHSPA1A transfection.

For bait protein production, cells were incubated for 72 hours following transfection before examination by light and fluorescence microscopy and harvesting of the culture supernatants. Culture supernatants were centrifuged at 2,000 x *g* for 10 minutes to pellet any detached cells. The supernatant was transferred to a fresh sterile tube and stored at 4 °C.

For pHSPA1A transfection, cells were incubated for 24 or 48 hours following transfection before being taken forward into an immunoprecipitation (section 2.4.7.2) or immunofluorescence (section 2.4.10) assay.

2.4 Protein techniques

2.4.1 ELISpot assay

Synthetic peptides (Pepscan, Lelystad, the Netherlands) were reconstituted to a stock concentration of 20 mM in dimethylsulfoxide (DMSO), then diluted to a working concentration of 1 mM in PBMC medium.

The assay was conducted using the Feline IFN-γ ELISpot Development and ELISpot Blue Colour Modules (Strep-AP and BCIP-NBT; R&D Systems, Abingdon, UK). Antibodies were reconstituted as per the manufacturer's instructions.

96-well MultiScreen-IP clear styrene sterile plates (Merck, Darmstadt, Germany) were coated with capture antibody. After blocking the plate with PBMC medium, extracted PBMCs were divided evenly across the appropriate wells and cultured for 24 hours (37 °C, 5% CO2) with either a positive control (concanavalin A type IV-S; Sigma-Aldrich), negative control (PBMC medium only) or peptide mix. The plate was then incubated overnight at 4 °C with the detection antibody. The plate was developed with Streptavidin-AP, followed by an incubation with BCIP/NBT Chromogen protected from light. Finally, the underdrain was removed and the plate washed thoroughly under tap water before drying completely. Washing was carried out between all steps using PBS and/or deionised water depending on the step.

Plates were photographed by an ELISpot Reader classic machine (Autoimmun Diagnostika, Strassberg, Germany) and spots counted by eye. Only medium and large spots with a fuzzy border were considered to be genuine spots. Number of spot forming cells (SFC) above background per 10^6 cells was calculated by subtracting mean spots under negative control conditions from mean spots under each peptide condition, then multiplying by 10^6/number of cells per well. Negative values (i.e. more spots in background than peptide wells) were displayed as zero. Individual spots were not counted in the positive control wells; instead a subjective assessment was made by comparison with the corresponding negative control.

2.4.2 FCoV antibody titre

Around 1 ml of whole feline blood was collected into tubes containing EDTA (whichever used normally by the veterinary practice in which the blood was being taken), stored at room temperature and processed as soon as practically possible after collection (usually within two days). Plasma, removed from the whole blood sample after allowing the blood cells to settle by gravity, was submitted to Veterinary Diagnostic Services, University of Glasgow, where an indirect

immunofluorescent antibody test was used to measure FCoV antibody titre. The antigen used in the test was Type 2 FCoV grown in FE-A cells.

2.4.3 Making a whole cell lysate

Cells grown in a cell culture flask were washed twice with cold PBS before addition of 500 µl 2x SDS-PAGE sample buffer. The cells were scraped into the sample buffer using a sterile cell scraper, transferred to a microcentrifuge tube and incubated at 95 °C for 5 minutes. The resulting lysate was chilled on ice then passed through a 25 G needle 10 times.

2.4.4 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

35 µl sample was incubated with 12.5 µl NuPAGE LDS Sample Buffer (Thermo-Fisher) and 2.5 µl β-mercaptoethanol at 70 °C for 10 minutes. A pre-cast NuPAGE 4-12% Bis-Tris mini gel (Thermo-Fisher) was assembled in an XCell SureLock Mini-Cell tank (Thermo-Fisher) in NuPAGE MOPS SDS running buffer (20x; Thermo-Fisher) as described by the manufacturer. 25 µl of protein sample was loaded into each well alongside 10 µl SeeBlue Plus2 Pre-stained Protein Standard (Thermo-Fisher). The gel was run at 200 V until the dye front reached the end of the gel (40-50 minutes).

2.4.5 Coomassie Brilliant Blue staining

Proteins separated by SDS-PAGE were visualised by soaking the gel in Coomassie Blue staining solution for at least 2 hours. The gel was then transferred to a destaining solution until the background was fully destained. Destaining solution was refreshed if necessary.

2.4.6 Western blot

Proteins separated by SDS-PAGE were transferred on to an Amersham Hybond polyvinylidene fluoride (PVDF) blotting membrane (GE Life Sciences, Pennsylvania, USA). A piece of membrane and 2 pieces of extra thick blot paper (Bio-Rad) were cut to the size of the gel. The extra thick blot paper and gel were soaked in transfer buffer whilst the membrane was first soaked in methanol and then soaked in transfer buffer. The transfer was carried out using a Trans-Blot Semi-Dry Transfer Cell (Bio-Rad, California, USA) run at 15 V for 50 minutes. The membrane containing transferred proteins was incubated with 5% (w/v) skimmed milk in PBS with 0.1% (v/v) Tween 20 (PBST) for 1 hour to block, then washed once in PBST. The membrane was incubated with primary antibody overnight at 4 °C. The membrane was then washed four times for 5 minutes per wash in PBST and incubated with secondary antibody for 1 hour. The membrane was washed four times for 5 minutes per wash in PBST and a chemiluminescent substrate (Amersham ECL Western Blotting Detection Reagent; GE Healthcare Life Sciences, Buckinghamshire, UK) was applied for 1 minute. A piece of Amersham

Hyperfilm x-ray film (GE Life Sciences) was exposed to the membrane for 5-30 seconds before the film was developed and fixed using a medical film processor (Photon Surgical Systems, Gloucester, UK).

2.4.7 Immunoprecipitation

2.4.7.1 Immunoprecipitation for proteomic analysis

CrFK cells were grown in cell line medium to 70-90% confluency in T225 flasks. Medium was poured off and the cell sheet was washed twice with PBS. The cells were scraped into cold PBS and centrifuged at 800 x *g* at 4 °C for 5 minutes.

Organoids embedded in Matrigel matrix were resuspended and transferred to cold PBS before centrifugation at 800 x *g* at 4 °C for 5 minutes. The supernatant was removed and the organoids were resuspended in cold PBS before centrifugation at 800 x *g* at 4 °C for 5 minutes.

In both cases, the supernatant was removed and the cells were resuspended in 1.5-3 ml cold cell surface lysis buffer. This was incubated on ice for 15 minutes with frequent agitation, before centrifugation at 14,000 x *g* at 4 °C for 10 minutes. The resulting lysate was divided into three aliquots and incubated with an equal volume of bait protein, plus extra n-Dodecyl ß-D-maltopyranoside to make a 0.6% (w/v) solution, at room temperature for 1 hour with mixing.

The resin from NAb Protein A Plus Spin Columns (Thermo-Fisher) was washed twice with PBS. Lysate-bait protein was incubated with the resin at 4 °C, with mixing, overnight, then the resin was put back on the columns. The resin was washed three times with PBS, then resuspended in 200 µl PBS prior to storage at -70 °C and analysis by the Proteomics Facility, University of Bristol.

At each stage, a 35 µl aliquot was taken for denaturation and analysis by SDS-PAGE and Coomassie Brilliant Blue staining (section 2.4.5) or western blotting (section 2.4.6).

2.4.7.2 Immunoprecipitation with pHSPA1A-transfected CrFK cells

pHSPA1A-transfected CrFK cells were grown in 6-well plates (section 2.3.17). Medium was poured off and the cell sheet was washed once with cold PBS. The cells were scraped into cold whole cell lysis buffer and incubated on ice for 30 minutes with frequent agitation, before centrifugation at 14,000 x *g* at 4 °C for 10 minutes.

An aliquot of the resulting lysate was taken for denaturation (14 parts lysate incubated with 1 part β-mercaptoethanol and 5 parts 4x SDS-PAGE sample buffer at 70°C for 10 minutes) and analysis by SDS-PAGE and western blotting (section 2.4.6). The remainder was divided into three aliquots and incubated with an equal volume of bait protein at room temperature for 1 hour with mixing.

The resin from NAb Protein A Plus Spin Columns (Thermo-Fisher) was washed twice with PBS. Lysate-bait protein was incubated with the resin at 4 °C, with mixing, overnight. The resin was washed three times with PBS, then proteins were eluted from the resin in elution buffer at 70°C for 10 minutes. Following centrifugation for 1 minute at 5000 x *g*, the supernatant was taken forward for analysis by SDS-PAGE and western blotting (section 2.4.6).

2.4.8 Proteomic analysis

2.4.8.1 TMT labelling

Pull-down samples were reduced (10 mM TCEP 55 °C, 1 hour), alkylated (18.75 mM iodoacetamide, room temperature, 30 minutes) and digested on the beads with trypsin (2.5 µg trypsin; 37 °C, overnight), then labelled with Tandem Mass Tag (TMT) six- or ten-plex reagents according to the manufacturer's protocol (Thermo-Fisher) and the labelled samples pooled.

The pooled sample was evaporated to dryness, resuspended in 5% (v/v) formic acid and then desalted using a SepPak cartridge according to the manufacturer's instructions (Waters, Milford, Massachusetts, USA). Eluate from the SepPak cartridge was again evaporated to dryness and resuspended in 1% (v/v) formic acid prior to analysis by nano-LC MSMS using an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo-Fisher).

2.4.8.2 Nano-LC Mass Spectrometry

The pooled TMT-labelled sample was fractionated using an Ultimate 3000 nano-LC system in line with an Orbitrap Fusion Tribrid mass spectrometer (Thermo-Fisher). In brief, peptides in 1% (v/v) formic acid were injected onto an Acclaim PepMap C18 nano-trap column (Thermo Scientific). After washing with 0.5% (v/v) acetonitrile, 0.1% (v/v) formic acid peptides were resolved on a 250 mm by 75 µm Acclaim PepMap C18 reverse phase analytical column (Thermo-Fisher) over a 150 minute organic gradient, using six gradient segments (5-9% solvent B over 2 minutes, 9-25% solvent B over 94 minutes, 25-60% solvent B over 23 minutes, 60-90% solvent B over 5 minutes, held at 90% solvent B for 5 minutes and then reduced to 1% solvent B over 2 minutes) with a flow rate of 300 nl min^{-1}. Solvent A was 0.1% (v/v) formic acid and Solvent B was aqueous 80% (v/v) acetonitrile in 0.1% (v/v) formic acid. Peptides were ionized by nano-electrospray ionization at 2.0 kV using a stainless-steel emitter with an internal diameter of 30 µm (Thermo-Fisher) and a capillary temperature of 275 °C.

All spectra were acquired using an Orbitrap Fusion Tribrid mass spectrometer controlled by Xcalibur 3.0 software (Thermo-Fisher) and operated in data-dependent acquisition mode using an SPS-MS3 workflow. FTMS1 spectra were collected at a resolution of 120,000, with an automatic gain control

(AGC) target of 200,000 and a max injection time of 50 milliseconds. Precursors were filtered with an intensity threshold of 5,000, according to charge state (to include charge states 2-7) and with monoisotopic peak detection set to peptide. Previously interrogated precursors were excluded using a dynamic window (60 s +/-10 ppm). The MS2 precursors were isolated with a quadrupole mass filter set to a width of 1.2 m/z. ITMS2 spectra were collected with an AGC target of 10,000, max injection time of 70 milliseconds and CID collision energy of 35%.

For FTMS3 analysis, the Orbitrap was operated at 30,000 (three-plex) or 50,000 (nine-plex) resolution with an AGC target of 50,000 and a max injection time of 105 milliseconds. Precursors were fragmented by high energy collision dissociation (HCD) at a normalised collision energy of 55% (three-plex) or 60% (nine-plex) to ensure maximal TMT reporter ion yield. Synchronous Precursor Selection (SPS) was enabled to include up to 5 MS2 fragment ions in the FTMS3 scan.

2.4.9 Flow cytometry

CrFK cells were grown in cell line medium to 70-90% confluency in T75 flasks prior to detachment with ACCUTASE™ Cell detachment solution (Stemcell) as per the manufacturer's instructions and prepared at 200,000 live cells in 1 ml FACS buffer (2% (v/v) FBS in PBS). PBMC were prepared from fresh or frozen at ≥ 2 x 10^5 live cells in 1 ml FACS buffer, depending on how many cells were available from that individual. Each sample was centrifuged at 300 x *g* for 3 minutes, the supernatant was poured off and the cells resuspended in 100 µl culture supernatant containing bait protein (1 in 2 dilution), human IgG Fc fragment (1 in 500 dilution; Merck) or FACS buffer. This was incubated for 20 minutes at room temperature. 1 ml FACS buffer was added to each sample and, after mixing, all were centrifuged at 300 x *g* for 3 minutes. The supernatant was poured off and the cells resuspended in 100 µl primary antibody for 15 minutes at room temperature. 1 ml FACS buffer was added to each sample and, after mixing, all were centrifuged at 300 x *g* for 3 minutes. The supernatant was poured off and the cells were resuspended in 100 µl FACS buffer and stored on ice prior to analysis by the Flow Cytometry Facility, University of Bristol, using an LSR II (BD Biosciences, San Jose, California) flow cytometer.

Data analysis was carried out using FlowJo software v10.3.

2.4.10 Immunofluorescence assay

Cell lines were cultured on glass coverslips until 70-100% confluent and PBMC were applied to glass slides by cytocentrifugation. Samples were fixed for 5 minutes using cold 4% (v/v) paraformaldehyde in PBS, permeabilized for 5 minutes using 0.1% (v/v) Triton X-100 in PBS then blocked for 1 hour using 10% (v/v) FBS in PBS. The fixed cells were incubated with culture supernatant containing bait

protein (1 in 2 dilution in 10% (v/v) FBS in PBS), human IgG Fc fragment (1 in 500 dilution in 10% (v/v) FBS in PBS) or primary antibody for 1 hour at room temperature or overnight at 4 °C.

Organoids embedded in Matrigel matrix were washed twice with PBS then suspended in PBS. The organoids were vigorously pipetted and passed through a 27 G needle before being transferred to microcentrifuge tubes and centrifuged at 200 x *g* for 5 minutes. The supernatant was removed and the organoids were resuspended in conditioned medium, then an equal volume of bait protein (1 in 2 final dilution in 10% (v/v) FBS in PBS) or human IgG Fc fragment (1 in 500 final dilution in 10% (v/v) FBS in PBS) was added. This was incubated at 4 °C for 1 hour with frequent agitation. Following centrifugation at 200 x *g* for 5 minutes and resuspension in PBS, the organoids were applied to glass slides by cytocentrifugation. Samples were fixed for 5 minutes using cold 4% (v/v) paraformaldehyde in PBS, permeabilized for 5 minutes using 0.1% (v/v) Triton X-100 (Sigma-Aldrich) in PBS then blocked for 1 hour using 10% (v/v) FBS in PBS.

In all cases, cells were then incubated with secondary antibody for 1 hour at room temperature. The cells were washed with PBS before and after all steps. The coverslips were mounted cell side down on glass microscope slides, or glass coverslips were applied to the cytocentrifugation cell spots, using VECTASHIELD Antifade Mounting Medium (Vector Laboratories) and clear nail varnish to seal the edges.

Slides were viewed and images taken in the Wolfson Bioimaging Facility, University of Bristol, using a Leica DM I6000 inverted epifluorescence microscope, a Leica SP5 confocal laser scanning microscope and Leica LAS-X acquisition software (Leica Microsystems, Milton Keynes, UK). Images were analysed post-acquisition using Fiji software.

2.5 Statistical analysis

In all cases, statistical significance was assigned at a level of $p<0.05$. All p values are given to three significant figures.

2.5.1 ELISpot

ELISpot data were compiled into a spreadsheet (Excel 2013, Microsoft) and exported into R Studio, running R statistical package v. 3.2.2. Bar and pie charts were generated in Excel, while all other graphs were produced in R. A Fisher's exact test was used to determine the significance of the difference in seroprevalence between 'Breeder' and 'Non-breeder' cats. The ELISpot data were evaluated for normal distribution using a histogram, q-q plot and Shapiro-Wilk test. A Kruskal-Wallis and/or Mann-Whitney test was used to determine the significance of any differences observed between peptide conditions on ELISpot. A Spearman's rank-order correlation coefficient was used to determine the significance of the relationship between plasma antibody titre and response to peptides.

2.5.2 Proteomic data analysis

The raw data files were processed and quantified using Proteome Discoverer software v2.1 (Thermo-Fisher) and searched against the UniProt *Felis catus* database using the SEQUEST algorithm. Peptide precursor mass tolerance was set at 10 ppm, and MS/MS tolerance was set at 0.6 Da. Search criteria included oxidation of methionine (+15.9949) as a variable modification and carbamidomethylation of cysteine (+57.0214) and the addition of the TMT mass tag (+229.163) to peptide N-termini and lysine as fixed modifications. Searches were performed with full tryptic digestion and a maximum of one (three-plex) or two (nine-plex) missed cleavages were allowed. The reverse database search option was enabled, and all data was filtered to satisfy false discovery rate (FDR) of 5%.

Data was filtered to exclude known contaminants. The raw abundances were log2 transformed, then sorted based on the fold change in their abundance compared with the negative control immunoprecipitation. A paired t-test was carried out to determine whether a given protein was significantly enriched across all replicates of the proteomics experiment.

3 Establishing a reverse genetic system for Type 1 FCoV

3.1 Introduction

This chapter will describe the construction of full-length recombinant viruses based on a field strain of Type 1 FCoV, which represents a step towards establishing a reverse genetic system for Type 1 FCoV.

The first FCoV reverse genetic system used targeted recombination to establish a Type 2 FCoV strain 79-1146 mutant, incorporating the MHV S protein ectodomain in an intermediate step that allowed recombinant viruses to be selected for (Haijema *et al.*, 2003). Though an important breakthrough in FCoV reverse genetics that was used to generate candidate FCoV vaccine strains (Haijema *et al.*, 2004), targeted recombination requires the recombinant virus to grow in cells, therefore does not allow study of lethal mutations or viruses that cannot be propagated in cell culture. Since selection of recombinant viruses is usually based on cell tropism, targeted recombination systems do not easily allow manipulation of the genome 5′ to the S gene (Almazan *et al.*, 2014).

A Type 1 FCoV reverse genetic system was first developed by Tekes *et al.* (2008) by cloning a cDNA copy of the FCoV strain 'Black' genome into a vaccinia virus vector. Using this system, the accessory gene cluster ORF3a-c was replaced with reporter genes, resulting in a recombinant virus that replicated as efficiently as the parental virus in FCWF cells. This reverse genetic system has been used to develop and characterise a recombinant Type 2 FCoV based on the strain '79-1146' and a chimeric virus based on Type 1 FCoV 'Black' but bearing the Type 2 FCoV '79-1146' S gene (Tekes *et al.*, 2010, Tekes *et al.*, 2012, Thiel *et al.*, 2014). Though these studies have vastly increased our understanding of the importance of the S gene in tropism and pathogenesis, they are limited in that 'Black' is highly laboratory-adapted and cannot be said to represent field strains of Type 1 FCoV (Pedersen, 2009). At the start of this project, the 'Black' system was the only Type 1 FCoV reverse genetic system that had been reported. It was therefore apparent that a reverse genetic system based on a field strain of Type 1 FCoV would be an invaluable tool in FCoV research, as it would allow the effect of viral mutations on the viral phenotype to be investigated using a clinically relevant strain. A reverse genetic system based on a field strain of Type 1 FCoV has since been established by another group (Ehmann *et al.*, 2018), and this will be discussed further in the discussion section of this chapter (section 3.3).

Due to their large RNA genomes, it was thought for a time that building full-length infectious cDNA clones of coronaviruses may not be possible. However, three strategies have enabled the development of coronavirus full-length infectious clones: vaccinia virus vectors, BACs and *in vitro*

ligation (Almazan *et al.*, 2014). The method used for building infectious clones in this project will be *in vitro* ligation, but sequences flanking the genome will be incorporated to allow the cDNA clone to be ligated into a vaccinia virus vector if necessary. The infectious clones in this project are based on the Type 1 FCoV 80F strain, which was isolated from the faeces of a naturally infected cat without FIP (Lewis *et al.*, 2015). This strain was chosen because it represents a typical non-virulent Type 1 FCoV field strain, therefore is clinically relevant. As well as a Type 1 FCoV 80F infectious clone ('T1'), a construct will be synthesised with a Type 1 FCoV 80F backbone but a Type 2 S gene ('T2S'). Such chimeric viruses have been shown to infect cells otherwise uninfectable by Type 1 FCoV and display similar *in vitro* growth kinetics to the Type 2 FCoV from which the spike gene was taken (Tekes *et al.*, 2010). Until a cell type in which T1 can be propagated is available to us, T2S would be a valuable construct for demonstrating the viability of the reverse genetic system and exploring non-spike mutations in the FCoV genome.

Since the N protein is required for efficient viral RNA synthesis, co-transfecting cells with full-length construct and FCoV N gene RNA has been shown to promote infectivity of the infectious clones (Yount *et al.*, 2000). It is thought that the N protein promotes infectivity by interacting directly with elements of the RTC (Baric *et al.*, 1988, Stohlman *et al.*, 1988, Hurst *et al.*, 2010). An alternative to co-transfecting with N gene transcripts is to transfect infectious clone RNA into cells stably expressing N protein (Tekes *et al.*, 2008), but the former method was utilised in this project because this is the method utilised by other groups using *in vitro* ligation.

In addition to the two infectious clones T1 and T2S, this project aimed to synthesise an FCoV replicon; something which has not been done previously. Replicons are self-replicating viral genomes from which some or all of the structural genes have been deleted and, in some cases, reporter genes have been added (Zimmer, 2010). Replicons contain all the genetic material necessary for viral RNA synthesis, but they cannot produce infectious progeny. In this way, replicons provide a safe way of studying the non-structural components of pathogenic viruses. Additionally, stable cell lines containing non-cytopathic replicons have also been produced, which provide a fast and simple way of analysing the effect of viral infection on the host cell and screening antiviral drugs (Almazan *et al.*, 2014). The same TGEV replicon construct was cytopathic in some cell lines but non-cytopathic in others (Almazan *et al.*, 2004), so selection of a suitable cell line is an important consideration in this project.

The first coronavirus replicon-like construct comprised the 5′ proximal ~20 kb of the HCoV 229E genome (encompassing the 5′ UTR and ORF1a-b gene), a GFP coding sequence, the 3′ UTR of the HCoV 229E genome and a synthetic poly-A tail. A small proportion of cells transfected with this

construct expressed GFP, and RNA extracted from these cells showed the typical organisation of a coronavirus subgenomic mRNA. These results demonstrated that the coronavirus replicase proteins suffice for subgenomic mRNA synthesis and translation, but did not show that the construct had replicated (Thiel *et al.*, 2001). Almazan *et al.* (2004), in establishing a replicon for TGEV, found that the N gene was the only structural gene necessary for viral genome replication. The TGEV replicon construct included an ORF7 accessory gene, but other studies have shown that the accessory genes are not necessary for virus replication in cell culture (Haijema *et al.*, 2004).

In this project, only the S and M genes were deleted from the replicon construct. This is because, although not necessary for viral genome replication, the E and accessory genes do not appear to be deleterious to the host cells (Ortego *et al.*, 2002) and their inclusion enhances virus replication in some scenarios (Dedeurwaerder *et al.*, 2013). Additionally, a puromycin selection marker and two reporter genes, GFP and *Renilla* luciferase, were included in order to select for cells containing the replicon and facilitate downstream analyses respectively.

3.1.1 Aims

The aims of this part of the project were to:

- Develop full-length cDNA constructs representing Type 1 FCoV (T1), Type 1 FCoV with a T2 spike gene (T2S) and an FCoV replicon using an *in vitro* ligation method
- Produce full-length *in vitro* RNA transcripts from these constructs
- Transfect mammalian cells with the full-length *in vitro* RNA transcripts in order to recover infectious virus (or replicon-expressing cells).

3.2 Results

3.2.1 Design and synthesis of cDNA fragments

Twenty-five cDNA fragments spanning the Type 1 FCoV strain 80F genome, the Type 2 FCoV strain 79-1146 S gene, a GFP puro resistance coding sequence and a *Renilla* luciferase coding sequence were designed. In various combinations, these fragments comprise three full-length constructs: Type 1 FCoV 80F ('T1'), Type 1 FCoV 80F backbone with a Type 2 FCoV 79-1146 spike gene ('T2S') and a Type 1 FCoV 80F with the S and M genes replaced with the GFP puro and *Renilla* luciferase coding sequences ('replicon'). All constructs were designed to incorporate a T7 promoter at the 5' end of the viral cDNA clone to enable them to function as templates for *in vitro* transcription of RNA. The 3' end of all constructs incorporated a poly-A tail and hepatitis delta virus (HDV) ribozyme to aid RNA processing in the host cell, but these were separated by a unique restriction endonuclease site to give the option of not using the ribozyme. To construct cDNA clones corresponding to the full-length genomes, a two-step strategy was used. In the first step, sets of three fragments were assembled using a commercially available NEBuilder HiFi DNA Assembly kit into 'units' (described in detail in section 3.2.3), before being ligated into full-length constructs in the second step (Figure 3.1).

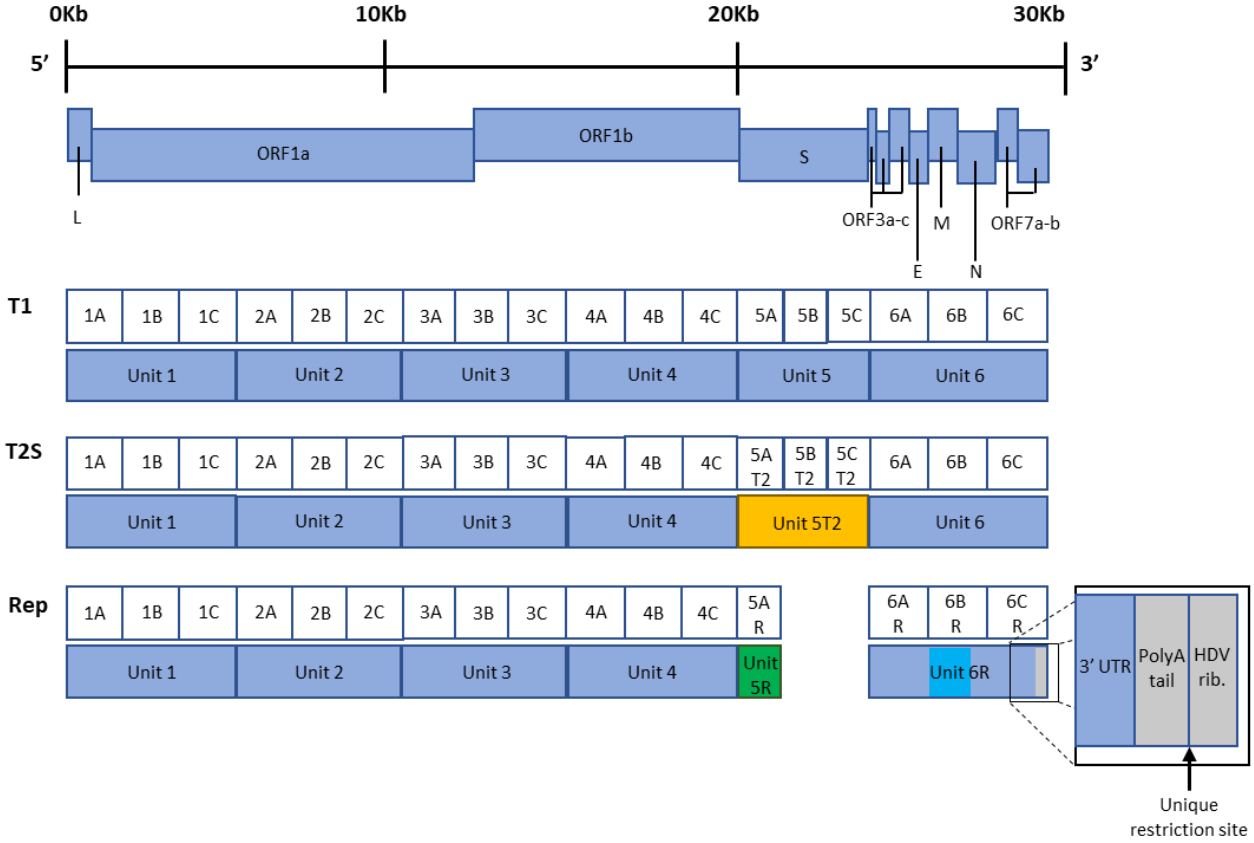

Figure 3.1. Schematic to show the fragments and units comprising each full-length construct. Fragments are shown as white boxes. Each set of three fragments – 'A', B' and 'C' – were assembled into a unit, except for 5AR which constituted unit 5R alone. The units were then ligated, in various combinations, into three full-length constructs: T1, T2S and replicon (Rep). The genome shown at the top of the figure is that of Type 1 FCoV strain 80F, on which the constructs are based. Units representing 80F are shown in light blue, the unit representing the Type 2 FCoV strain 79-1146 S gene is shown in yellow, a GFP puro coding sequence is shown in green, and *Renilla* luciferase coding sequence is shown in bright blue. A closer view of the 3' end of the replicon reveals a poly-A tail and HDV ribozyme, shown in grey, after the 3' UTR, separated by a unique restriction endonuclease site. This is the same for all constructs.

The cDNA fragment sequences were submitted to Twist Bioscience for synthesis and cloning into the 2,221 bp pTwist 31 pMB1-Amp-UCS1.2 (pTwist) vector. Twenty-one of the 25 cDNA fragments arrived as plasmids. The remainder could not be amplified in bacterial culture so arrived as synthetic DNA fragments. Table 3.1 summarises the cDNA fragments and Appendix A shows the full sequences.

Table 3.1. The cDNA fragments synthesised for construction of recombinant viruses. Twenty-five fragments were synthesised in total, and of these 21 arrived as plasmids. Four could not be cloned into plasmids so arrived as synthetic DNA fragments. The 'unit' and 'construct' columns indicate which unit each fragment belongs to, and which construct each unit belongs to.

Fragment	Description	Plasmid?	Unit	Construct
1A	5' UTR, Type 1 FCoV 80F ORF1a/b gene	Yes		
1B	Type 1 FCoV 80F ORF1a/b gene	Yes		
1C	Type 1 FCoV 80F ORF1a/b gene	Yes	1 (5,114 bp)	All
2A	Type 1 FCoV 80F ORF1a/b gene	Yes		
2B	Type 1 FCoV 80F ORF1a/b gene	Yes		
2C	Type 1 FCoV 80F ORF1a/b gene	Yes	2 (5,112 bp)	All
3A	Type 1 FCoV 80F ORF1a/b gene	Yes		
3B	Type 1 FCoV 80F ORF1a/b gene	Yes		
3C	Type 1 FCoV 80F ORF1a/b gene	Yes	3 (5,112 bp)	All
4A	Type 1 FCoV 80F ORF1a/b gene	Yes		
4B	Type 1 FCoV 80F ORF1a/b gene	Yes		
4C	Type 1 FCoV 80F ORF1a/b gene	Yes	4 (5,114 bp)	All
5A	Type 1 FCoV 80F spike gene	No		
5B	Type 1 FCoV 80F spike gene	Yes		
5C	Type 1 FCoV 80F spike gene	Yes	5 (4,415 bp)	T1
5AR	Green fluorescent protein/puromycin resistance coding sequence	Yes	5R (1,372 bp)	Replicon
5AT2	Type 2 FCoV 79-1146 spike gene	No		
5BT2	Type 2 FCoV 79-1146 spike gene	Yes		
5CT2	Type 2 FCoV 79-1146 spike gene	Yes	5T2 (4,373 bp)	T2S
6A	Type 1 FCoV 80F ORF3a-c, envelope and membrane genes	Yes		
6B	Type 1 FCoV 80F membrane and nucleocapsid genes	Yes		
6C	Type 1 FCoV 80F nucleocapsid and ORF7a/b genes, 3' UTR, poly-A tail, HDV ribozyme	No	6 (4,547 bp)	T1, T2S
6AR	Type 1 FCoV 80F ORF3a-c and envelope genes, *Renilla* luciferase coding sequence	Yes		
6BR	*Renilla* luciferase coding sequence, Type 1 FCoV 80F nucleocapsid gene	Yes		
6CR	Type 1 FCoV 80F nucleocapsid and ORF7a/b genes, 3' UTR, poly-A tail, HDV ribozyme	No	6R (4,691 bp)	Replicon

3.2.2 Amplifying fragments

Bacteria (*E. coli*) were transformed with all 21 plasmids to amplify the plasmids for downstream applications. Following transformation, at least three bacterial colonies were selected per plasmid construct and the bacteria were cultured. Plasmid DNA was then extracted using a miniprep or midiprep procedure, and a restriction digest was carried out to check the size of the digestion products.

For fragment 1B, a construct of the correct size was identified and taken forward for assembly of unit 1 (1B8; Figure 3.2). However, following multiple failed attempts to ligate unit 1, plasmid 1B8 was Sanger sequenced and found to be aberrant in that the sequence was not what was expected. Eventually a fragment 1B with the correct sequence was obtained by PCR amplification of the fragment from the plasmid provided by Twist Bioscience, using primers that sit within the pTwist vector (Twist For and Twist Rev; Appendix A). The difficulties experienced during the construction of unit 1 are described in detail in section 3.2.3.1. Amplified plasmid DNA of the correct size was obtained easily for the other fragments available as plasmids (Figure 3.3), though 2A and 4B showed a degree of instability in bacterial culture with multiple colonies screened before a construct of the correct size was identified.

Fragments were liberated from their plasmids or flanking sequences through restriction digestion. The digestion products were separated by agarose gel electrophoresis and DNA bands of the correct size were excised from the gel and extracted.

Those fragments not available as plasmids were not amplified prior to downstream applications.

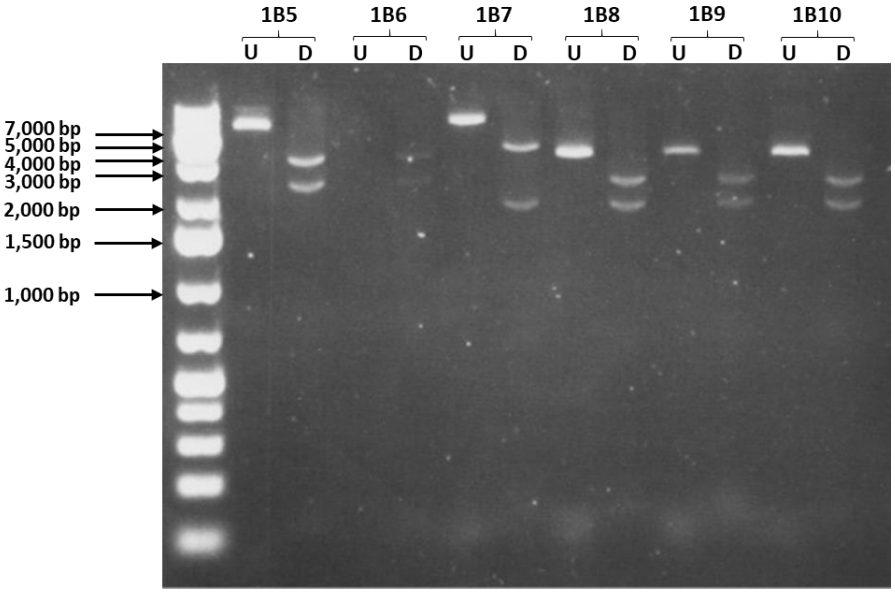

Figure 3.2. Analysis of fragment 1B by agarose gel electrophoresis. Six clones containing fragment 1B (1B5-10) were amplified in bacterial culture, then the plasmid DNA was extracted and digested with appropriate restriction enzymes to liberate the fragments from their plasmids. The digestion products (D) were run alongside an aliquot of undigested plasmid (U) on a 1% agarose gel at 100 V for 40 minutes. The DNA bands were visualised under a UV transilluminator. The positions of relevant DNA mass markers are shown in bp.

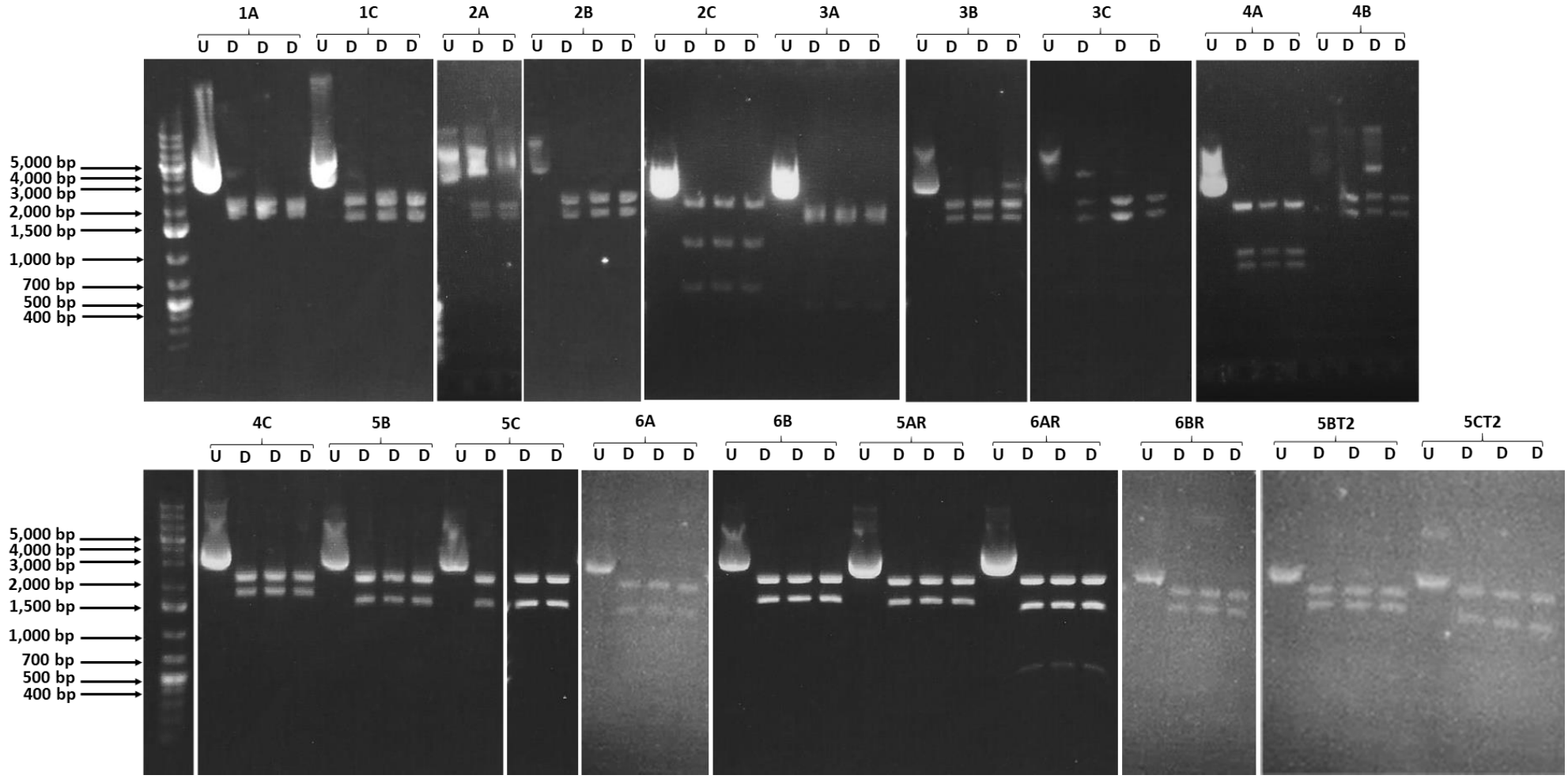

Figure 3.3. Analysis of all fragments available as plasmids by agarose gel electrophoresis. Two to three clones containing fragments 1A, 1C, 2A, 2B, 2C, 3A, 3B, 3C, 4A, 4B, 4C, 5B, 5C, 6A, 6B, 5AR, 6AR, 6BR, 5BT2 and 5CT2 were amplified in bacterial culture, then the plasmid DNA was extracted and digested with appropriate restriction enzymes to liberate the fragments from their plasmids. The digestion products (D) were run alongside an aliquot of undigested plasmid (U) on a 1% agarose gel at 100 V for 40 minutes. The DNA bands were visualised under a UV transilluminator. The positions of relevant DNA mass markers are shown in bp. Fragments 5A, 6C, 6CR and 5AT2 are not shown because these fragments were not available as plasmids and fragment 1B is shown in Figure 3.2.

3.2.3 Assembling units

In the first step of full-length cDNA clone production, units were assembled from their constituent fragments using the NEBuilder HiFi DNA assembly method. Fragments were designed to overlap by 25 bp, which enables an exonuclease to chew back the 5’ ends of DNA strands, allowing the complementary 3’ overhangs of overlapping sections to match and form a seamless join (Figure 3.4).

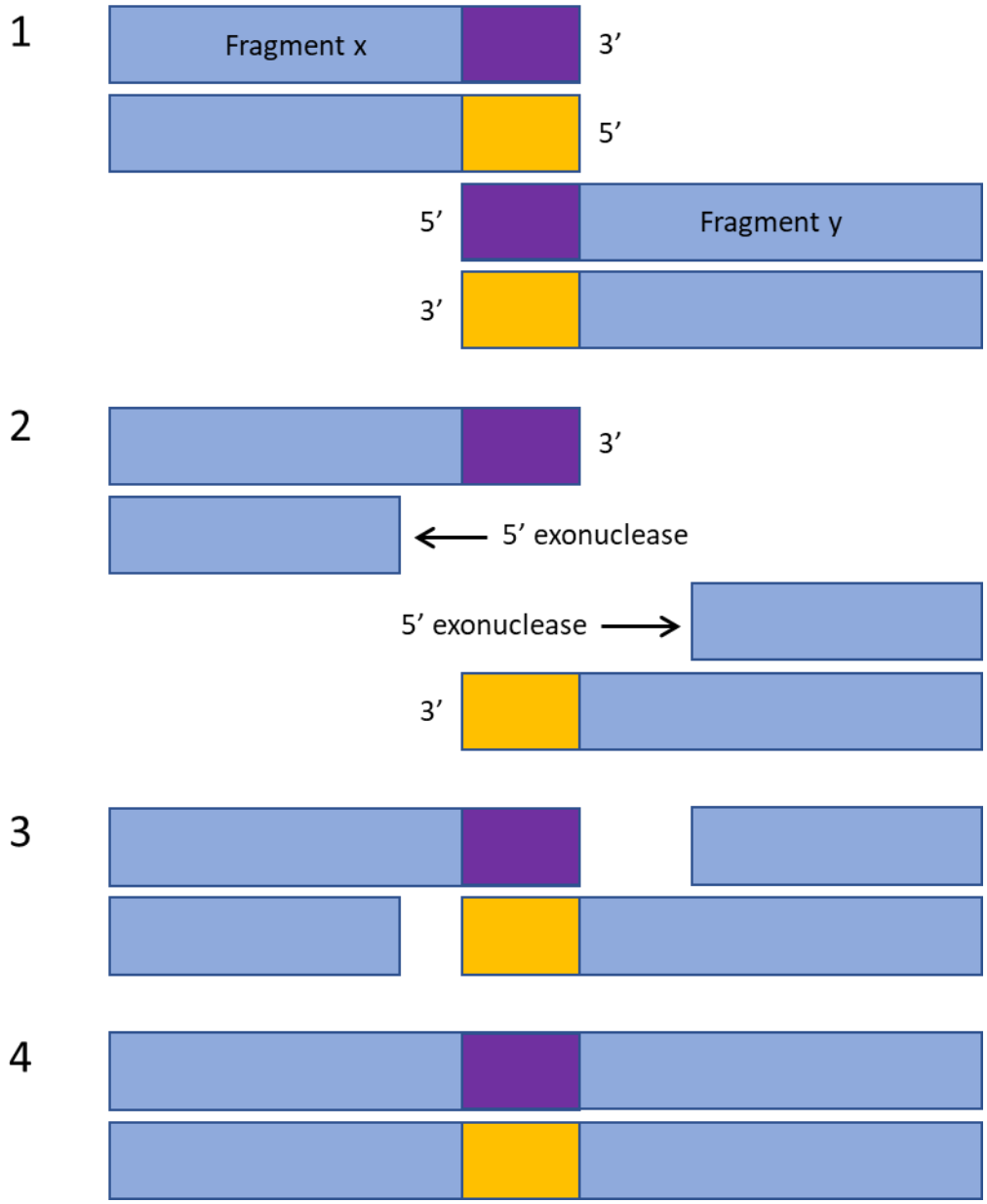

Figure 3.4. Schematic to illustrate the NEBuilder HiFi DNA assembly method.
1: Two DNA fragments (shown as double-stranded light blue bars) are designed to overlap by 25 bp. Overlapping sequences are shown in yellow or purple, depending on which DNA strand is being considered.
2: A 5’ exonuclease chews back the 5’ ends of DNA strands, revealing the complementary sequences on the 3’ ends.
3: The complementary sequences match and anneal.
4: A DNA polymerase fills in the gaps on the 5’ strands and a DNA ligase seals the nicks, creating a single piece of seamlessly ligated DNA.

The unit assembly strategy utilised in this project used restriction digestion to linearise the plasmid containing the ‘A’ fragment and liberate the ‘B’ and ‘C’ fragments from their plasmids. The ‘B’ and ‘C’ fragments were then assembled into the plasmid containing the ‘A’ fragment by NEBuilder HiFi DNA assembly. This resulted in a pTwist vector containing a complete unit (Figure 3.6A). This strategy was used to assemble units 1, 2, 3, 4, 6 and 6R.

Since ‘A’ fragments were not available as plasmids for units 5 and 5T2 and the ‘A’ fragments were obviously not stable in a high copy number plasmid vector, a different assembly strategy was adopted. Oligonucleotide adaptors (Appendix A) overlapping with the 5’ and 3’ ends of these units were designed and synthesised. Figure 3.5 shows an example of an adaptor used in this approach. The adaptors were designed with restriction endonuclease sites at each end that would allow them to be ligated into the multiple cloning site of the very low copy number vector pWSK29. The adaptors also incorporated a restriction endonuclease site in the middle *via* which the vector could be linearised, revealing ends complementary to the ends of the unit. The ‘A’, ‘B’ and ‘C’ fragments, liberated from the respective plasmids, were then assembled into the introduced adaptor sequence in the linearised pWSK29 vector by NEBuilder HiFi DNA assembly (Figure 3.6B).

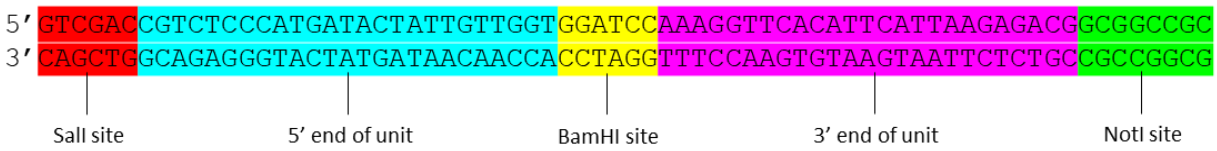

Figure 3.5. The oligonucleotide adaptor used to ligate unit 5 into a low copy number vector. The adaptor was flanked by restriction endonuclease sites (SalI and NotI) that would allow it to be ligated into low copy number vector pWSK29 *via* the vector’s multiple cloning site. In the middle of the adaptor, a BamHI restriction endonuclease site allowed the adaptor-containing vector to be linearised, revealing ends complementary to the 5’ and 3’ ends of unit 5 that could be used for NEBuilder HiFi DNA assembly.

Unit 5R consisted only of fragment 5AR so did not need to undergo further assembly.

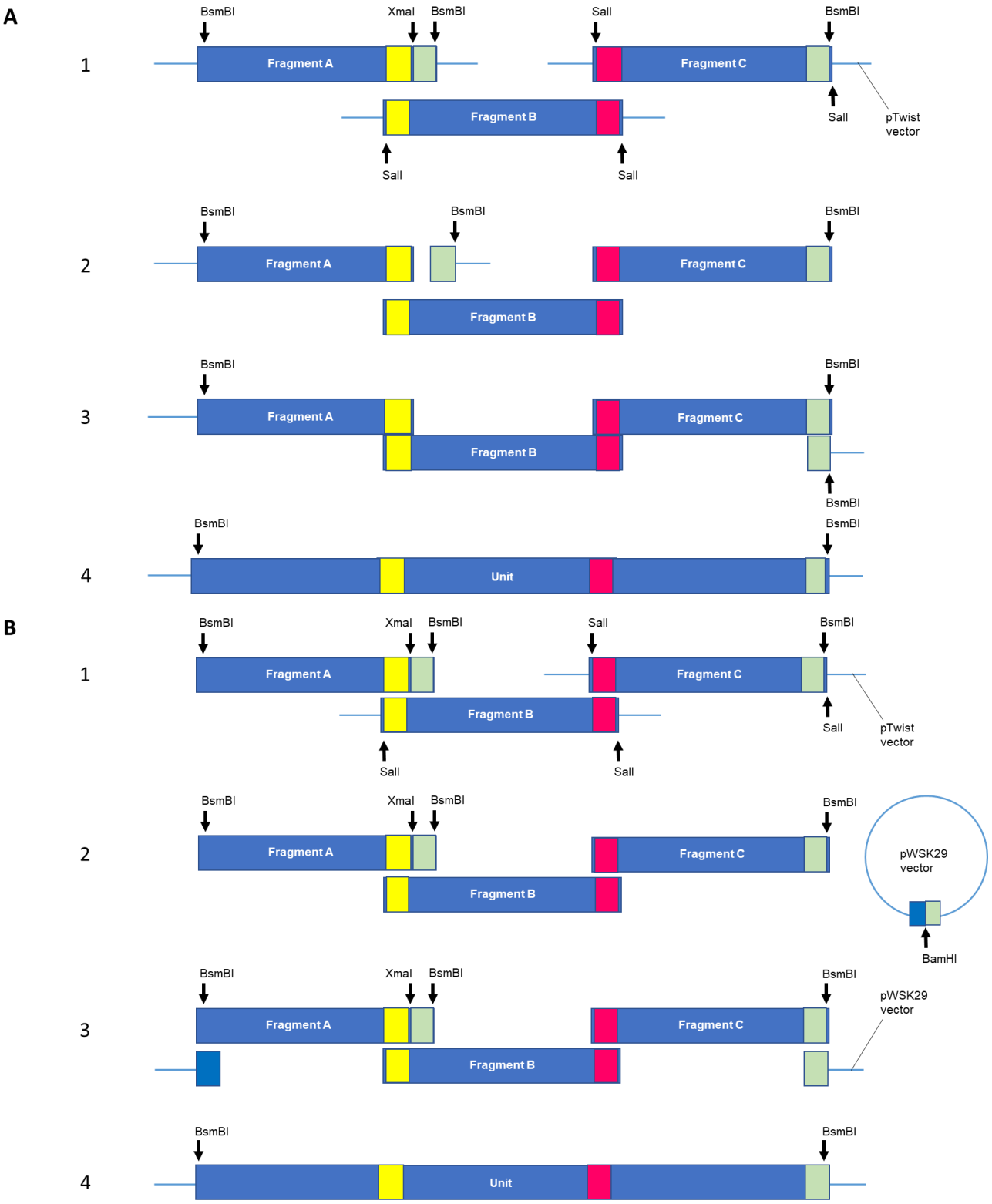

Figure 3.6. Schematic to show the assembly of an infectious clone 'unit' from three fragments. Fragments are shown as blue bars. Overlapping regions are shown are smaller yellow, green or red bars within the fragments. Blue lines indicate vectors. Heavy black arrows indicate restriction endonuclease sites. Units 1, 2, 3, 4, 6 and 6R were ligated using method A, whereas units 5 and 5T2 utilised method B.
A1: Fragments 'B' and 'C' are liberated from their plasmids using SalI, and the plasmid containing fragment 'A' is linearised using XmaI.
A2: Note that both ends of fragment 'A' remain attached to the vector.
A3: NEBuilder HiFi DNA assembly is used to seamlessly assemble the fragments together *via* their overlapping regions.
A4: The three fragments are assembled into a single unit, residing within the vector that originally contained just fragment 'A'.
B1: Fragments 'B' and 'C' are liberated from their plasmids using SalI. Fragment 'A' is not digested.
B2: pWSK29 is digested with BamHI, which cleaves the adaptor revealing two regions that overlap with either end of the unit.
B3: NEBuilder HiFi DNA assembly is used to seamlessly assemble the fragments together and into pWSK29 *via* their overlapping regions.
B4: The three fragments are assembled into a single unit, residing within pWSK29.

After unit assembly using the NEBuilder HiFi DNA assembly protocol, bacteria were transformed with the resulting unit-containing plasmids. At least three bacterial colonies were selected per construct and the bacteria were cultured. Plasmid DNA was extracted using a miniprep or midiprep procedure and the DNA underwent restriction digestion and agarose gel electrophoresis to check the size of the digestion products. An example is shown in Figure 3.7, where Cl5 is the correct clone of unit 3. A clone of the correct size was identified fairly rapidly for all units except unit 1.

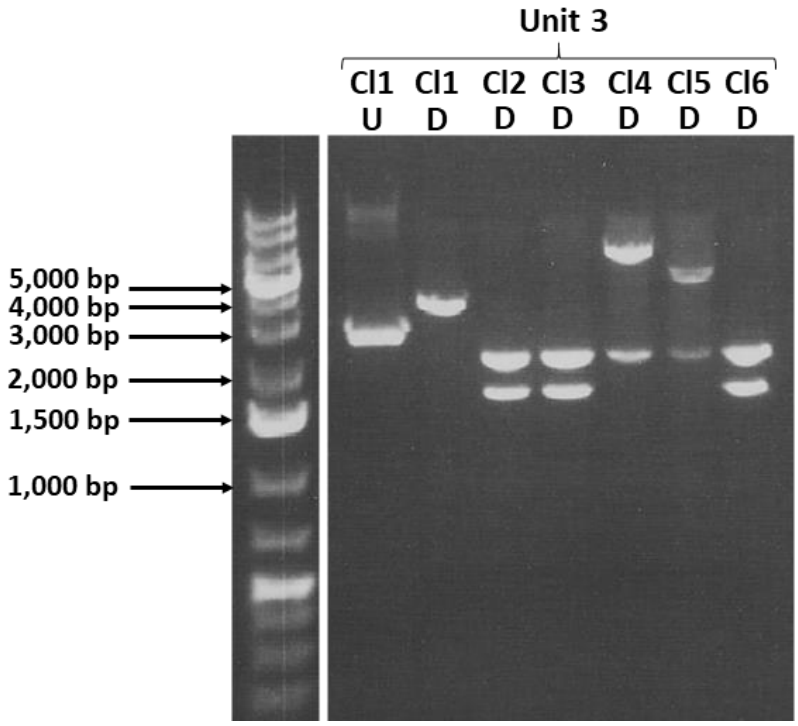

Figure 3.7. Analysis of unit 3 by agarose gel electrophoresis. Six clones (Cl1-6) of unit 3 were amplified in bacterial culture, then the plasmid DNA was extracted and digested with appropriate restriction enzymes to give a particular digestion pattern. The digestion products (D) were run alongside an aliquot of undigested plasmid (U) on a 0.7% agarose gel at 100 V for 40 minutes. The DNA bands were visualised under a UV transilluminator. The positions of relevant DNA mass markers are shown in bp.

3.2.3.1 Unit 1 assembly.

Despite selecting and culturing over 20 bacterial colonies for unit 1, a plasmid clone of the correct size was not identified. Many of the clones had different restriction digestion patterns and sizes to one another (Figure 3.8A), suggesting instability of the unit in bacterial culture. Culturing the bacteria at 28 °C was therefore attempted to increase plasmid stability, but this still did not yield clones of the correct size. It was then decided to use the same strategy utilised for units 5 and 5T2, where an oligonucleotide adaptor was designed, synthesised and ligated into pWSK29 (Figure 3.5B). Following another NEBuilder HiFi DNA assembly reaction to ligate fragments 1A, 1B and 1C into the modified pWSK29 plasmid, bacteria were transformed with the ligation reaction mix and over 20 colonies were selected for bacterial culture. Again, none of the plasmid clones were of the correct size, despite culturing the bacteria at 28 °C. An attempt was made to PCR amplify the unit from both ligation reactions, but an amplicon of the correct size was not obtained (data not shown).

Due to the difficulty experienced obtaining a correct clone of fragment 1B, it was suspected that this fragment may be causing an issue. Clones 8-10 of fragment 1B, which all appeared to be the correct size, were PCR amplified and Sanger sequenced. The sequence of 1B8, the clone that was being used for the unit 1 ligation, was found to be completely aberrant in that it was not what was expected. The sequences of 1B9 and 1B10 were correct, except for a single nucleotide insertion that would shift the reading frame and render them unusable. A sequence-perfect clone of fragment 1B was finally obtained through PCR amplification of the fragment from the plasmid provided by Twist Bioscience, using Twist For and Twist Rev primers (Appendix A). Unit 1 assembly was repeated with this new fragment, using the original strategy of ligating 1B and 1C into the pTwist vector containing 1A. Following bacterial transformation and culture, a single clone of the correct size (Cl51) was identified (Figure 3.8B).

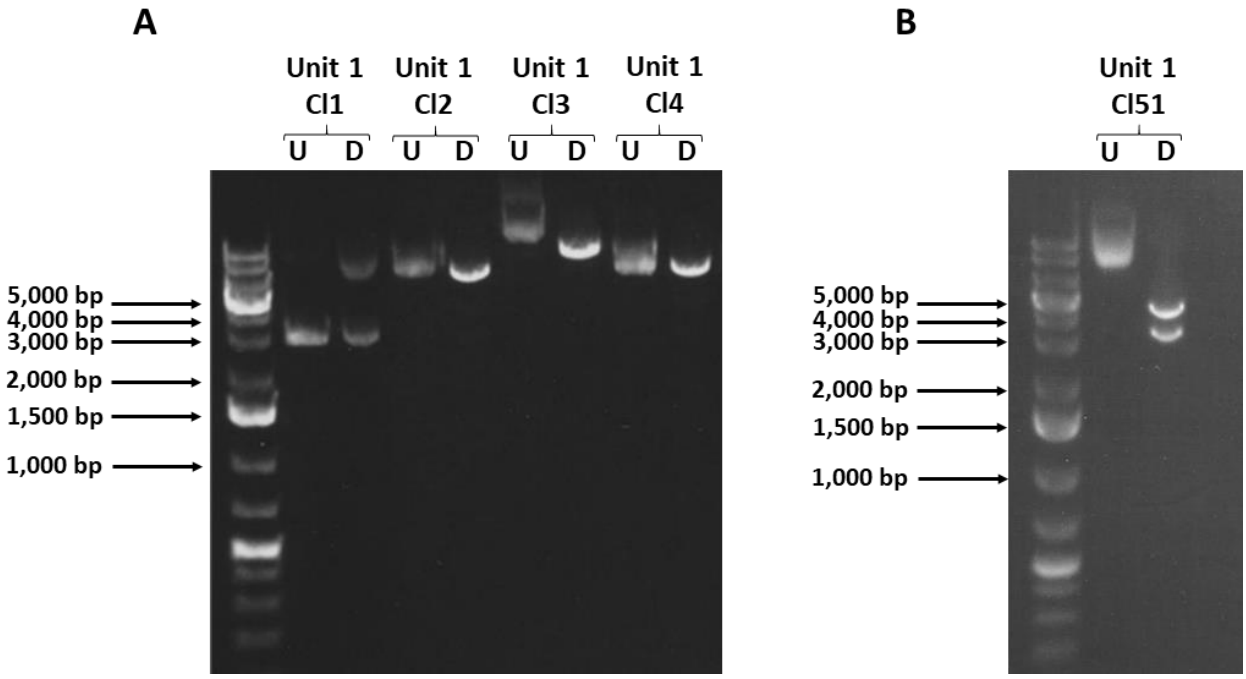

Figure 3.8. Analysis of unit 1 by agarose gel electrophoresis. Clones of unit 1 were amplified in bacterial culture, then the plasmid DNA was extracted and digested with appropriate restriction enzymes to give a particular digestion pattern. The digestion products (D) were run alongside an aliquot of undigested plasmid (U) on a 0.7% agarose gel at 100 V for 40 minutes. The DNA bands were visualised under a UV transilluminator. The positions of relevant DNA mass markers are shown in bp.
A: Four incorrect clones (Cl1-4), illustrating the heterogeneity of the clones.
B: A clone of the correct size (Cl51) was finally identified.

3.2.4 Obtaining sequence-perfect units

When a clone of the correct size was identified for a unit, plasmid DNA was submitted for Sanger sequencing following amplification by larger scale bacterial culture or PCR if necessary. See Appendix A for a full list of sequencing primers used in this project. All units were found to have the correct sequence except for units 1, 2 and 4.

Despite the effort that went in to producing a sequence-perfect fragment 1B, following amplification in bacterial culture, unit 1 was found to have three single nucleotide substitutions (T>C at unit 1

position 1796, T>C at unit 1 position 2225, A>G at unit 1 position 2487) within the fragment 1B region. The first two mutations are synonymous so would not be expected to influence the structure or function of the final infectious clone into which the unit would be incorporated, but the latter mutation results in an amino acid substitution of valine for isoleucine that would be situated in pp1a and pp1ab following translation of ORF1a-b. These mutations likely arose during bacterial culture following transformation of the ligated unit into bacteria, illustrating the instability of unit 1 (and particularly fragment 1B) in bacterial culture. It was decided to proceed with this imperfect unit due to time limitations for the project.

The clone of unit 2 submitted for sequencing (Cl8) was found to have a 410 bp deletion within the fragment 2A region. Five clones of unit 2 were therefore digested with appropriate restriction enzymes to assess whether the deletion was present (Figure 3.9). Two clones with the correct restriction digestion pattern (Cl4 and Cl7) were identified. Cl7 was found to be perfect on sequencing so was taken forward.

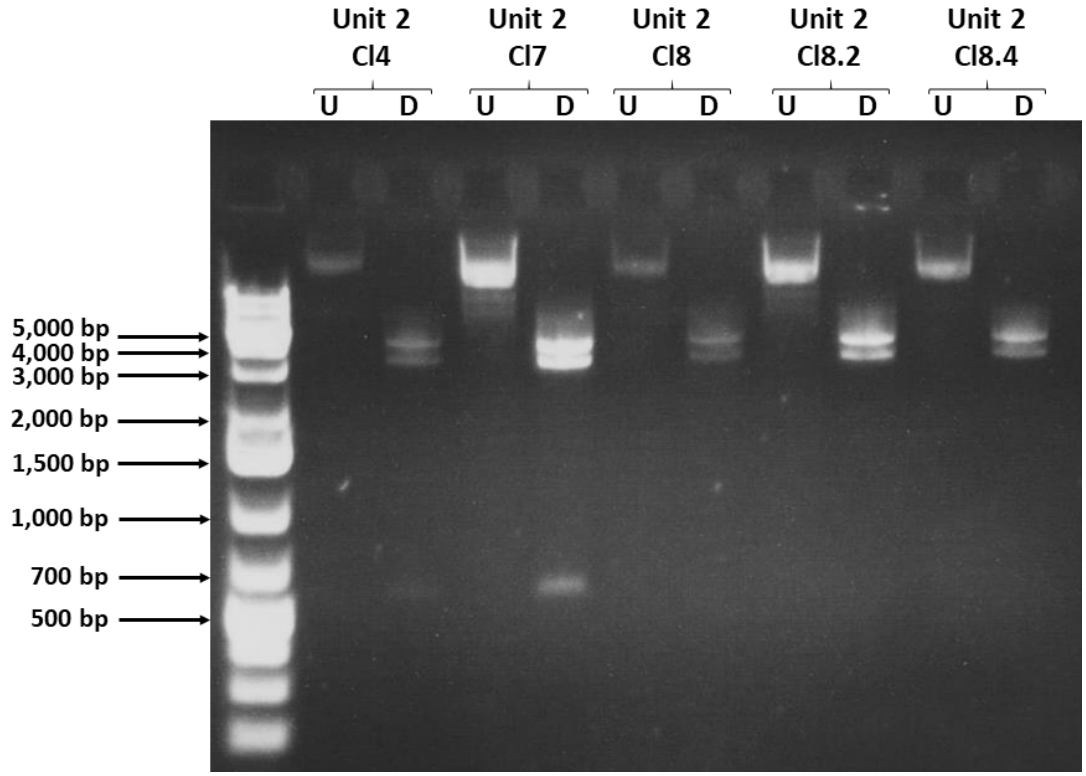

Figure 3.9. Analysis of unit 2 by agarose gel electrophoresis. Following identification of a 410 bp deletion in unit 2 clone 8 (Cl8) on sequencing, five clones of unit 2 were digested with appropriate restriction enzymes to assess whether the 410 bp deletion was present. The digestion products (D) were run alongside an aliquot of undigested plasmid (U) on a 0.7% agarose gel at 100 V for 40 minutes. The DNA bands were visualised under a UV transilluminator. The positions of relevant DNA mass markers are shown in bp.

3.2.4.1 Obtaining a sequence-perfect unit 4

The clone of unit 4 submitted for sequencing (Cl6) was found on sequencing to have a single nucleotide deletion in the region of complementarity between fragments 4B and 4C, which would shift the reading frame and render the unit unusable. Three more clones were sequenced over the region of the deletion, but all were found to have the same deletion. The NEBuilder HiFi DNA assembly ligation reaction was repeated for unit 4 as, being on the border of two fragments, it was unclear whether the deletion arose as a result of unit instability in bacterial culture or an issue with the assembly reaction. Bacteria were transformed with the new unit 4-containing plasmid, four clones were selected and the bacteria were cultured. Only one clone (Cl4N) had the correct restriction digestion pattern (Figure 3.10A), so this clone was sequenced over the region of the deletion. Unit 4 Cl4N was found to have the same deletion.

A 773 bp DNA fragment (Appendix A) was designed and synthesised to replace a region of unit 4 spanning the deletion. Naturally occurring restriction endonuclease sites either side of the deletion were identified and the fragment was flanked by these sites, allowing the mutated region to be excised and the fragment to be ligated into the unit in its place. The repaired unit was transformed into bacteria, eight clones were selected for bacterial culture and the plasmid DNA was extracted. On restriction digestion and agarose gel electrophoresis (Figure 3.10B), two clones (Cl6repair and Cl7repair) were the correct size. Unit 4 Cl7repair was sequenced but found to contain the same deletion, suggesting either that unit 4 is unstable in bacterial culture or the repair had failed.

An oligonucleotide adaptor (Appendix A) was designed and synthesised to overlap with the 5’ and 3’ ends of unit 4 (Figure 5), allowing the unit to be cloned into the low copy number vector pWSK29. Following another NEBuilder HiFi DNA assembly reaction to ligate fragments 4A, 4B and 4C into the low copy number vector, bacteria were transformed with the resulting plasmid and over 60 clones were selected for bacterial culture. None of the clones were of the correct size (Figure 3.10C), despite repeating the ligation twice using different insert to vector ratios and culturing the bacteria at 28 °C.

Eventually, a sequence-perfect unit 4 was obtained by PCR amplifying the unit from the first NEBuilder HiFi DNA assembly ligation reaction (Figure 3.10D) using Twist For and Twist Rev primers (Appendix A). Unit 4 only existed as a PCR amplicon and was not cloned into a vector.

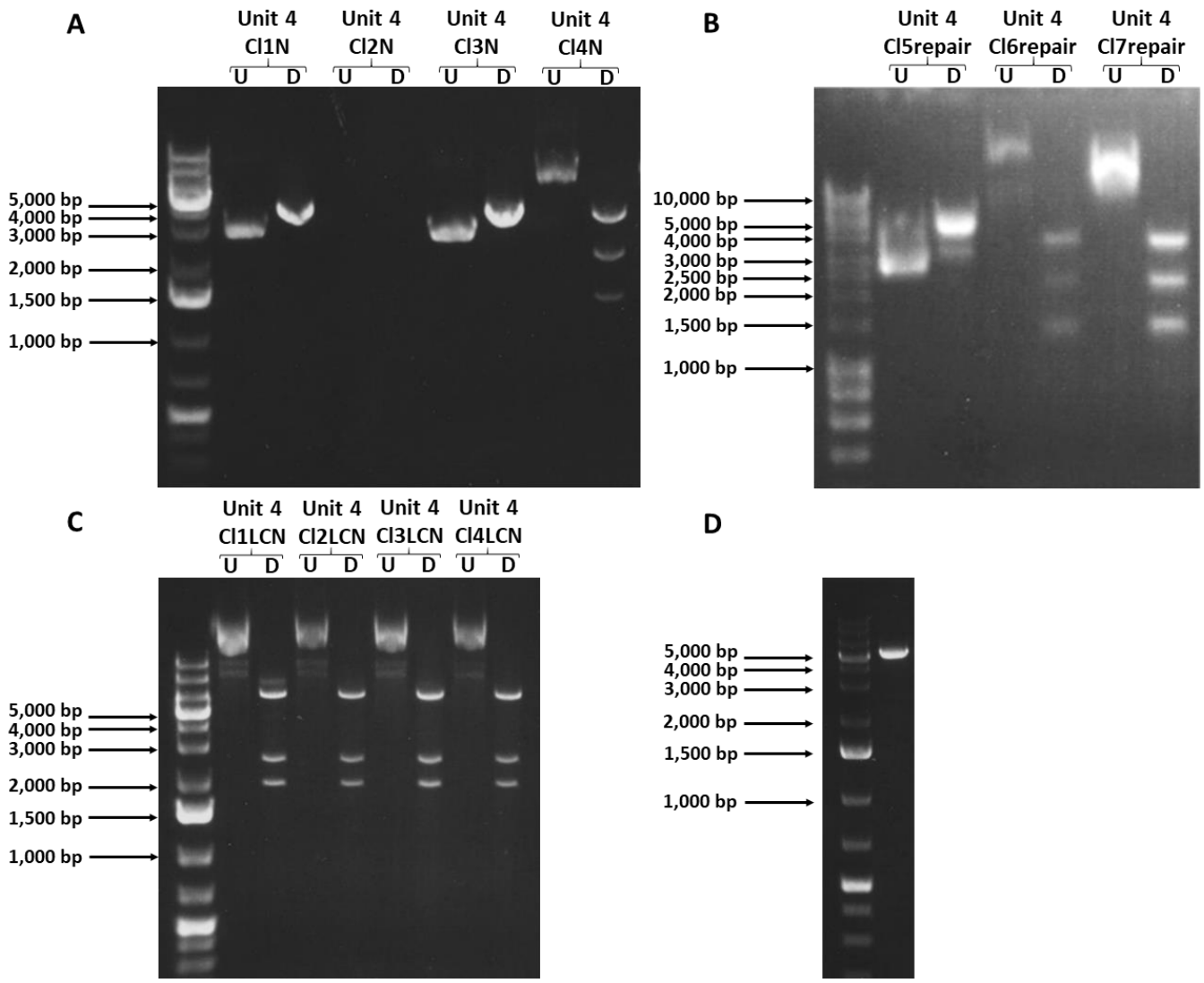

Figure 3.10. Analysis of unit 4 by agarose gel electrophoresis. All DNA (digestion products for A, B and C, and PCR product for D) was run on a 0.7% agarose gel at 100 V for 40 minutes. The DNA bands were visualised under a UV transilluminator. The positions of relevant DNA mass markers are shown in bp.
A: Following detection of a deletion in unit 4 Cl6, the unit 4 ligation reaction was repeated and transformed into bacteria. Four colonies (Cl1N-4N) were selected for bacterial culture. Plasmid DNA was extracted and digested with appropriate restriction enzymes to give a particular digestion pattern. The digestion products (D) were run alongside an aliquot of undigested plasmid (U).
B: The repaired unit 4 was transformed into bacteria and eight colonies were selected for bacterial culture. Plasmid DNA was extracted and digested with appropriate restriction enzymes to give a particular digestion pattern. The digestion products (D) were run alongside an aliquot of undigested plasmid (U). Three of the eight clones (Cl5repair-Cl7repair) are shown here.
C: Unit 4 was ligated into a low copy number (LCN) vector, the plasmid was transformed into bacteria and over 60 colonies were selected for bacterial culture. Plasmid DNA was extracted and digested with appropriate restriction enzymes to give a particular digestion pattern. The digestion products (D) were run alongside an aliquot of undigested plasmid (U). Four of the clones (Cl1LCN-Cl4LCN) are shown here.
D: Unit 4 was PCR amplified from the first ligation reaction using primers that sit within the pTwist vector.

3.2.5 Amplifying units

All units had to be amplified to produce sufficient DNA to be ligated into full-length constructs. Sequence verified plasmid clones of units 1, 3, 6, 5R and 6R were amplified by large scale bacterial culture and plasmid DNA was extracted using a midiprep or maxiprep procedure. This approach was attempted for units 5 and 5T2, but it was not possible to obtain enough plasmid DNA for unit extraction and ligation, as these units reside in a very low copy number vector. Units 5 and 5T2 were

therefore amplified by PCR from sequenced clones, using a proof-reading polymerase and primers that are complementary to the pWSK29 vector (Appendix A).

In an attempt to amplify unit 2 for downstream applications, unit 2 Cl7 plasmid DNA was transformed into bacteria then 10 clones were selected for large scale bacterial culture. However, none of these 10 clones had the correct restriction digestion pattern, suggesting that unit 2 is unstable in large scale bacterial culture. Unit 2 was therefore amplified by PCR from Cl7 using Twist For and Twist Rev primers that had been adapted to add flanking sequences on to unit 2 to enable it to be cloned into pWSK29 (Appendix A). However, unit 2 was not cloned into pWSK29 during this project.

Unit 4 had to be amplified by PCR since the only correct version of the unit did not reside in a vector. This was carried out using Twist For and Twist Rev primers (Appendix A).

All amplified units were sequenced again and found to be unchanged. See Appendix A for the final unit sequences.

3.2.6 Ligation of units to produce full-length cDNA clones

As all units were verified to be sequence-perfect (except for unit 1; section 3.2.4) and amplified to a sufficient level, the next step was to ligate the units to produce three full-length constructs: T1, T2S and replicon. Units were designed to incorporate unique BsmBI (type IIS restriction endonuclease) cleavage sites, thereby eventually allowing the units to be assembled seamlessly and directionally into a full-length construct (Figure 3.11). Units 2, 3, 4, 5, 5T2 and 5R had BsmBI sites at each end of the unit, flanking the genomic sequence, whereas unit 1 had a BsmBI site only at its 3' end and units 6 and 6R had BsmBI sites only at their 5' ends.

Unit 1 underwent restriction digestion to cut upstream of the unit and linearise the plasmid in which it sat. Units 6 and 6R underwent restriction digestion to cut downstream of the units and linearise the plasmids in which they sat. The digested plasmids were phosphatase-treated to prevent re-ligation. All units were then digested with BsmBI to liberate units from their plasmids (or flanking sequences in the case of PCR products) and produce uniquely compatible ends between units whilst maintaining the viral genome sequence. The digestion products were separated by agarose gel electrophoresis (Figure 3.12A) and DNA bands corresponding to the liberated units were excised from the gel and extracted (Figure 3.12B). All units were of the correct size (see Table 3.1 for expected sizes) and adequate purity.

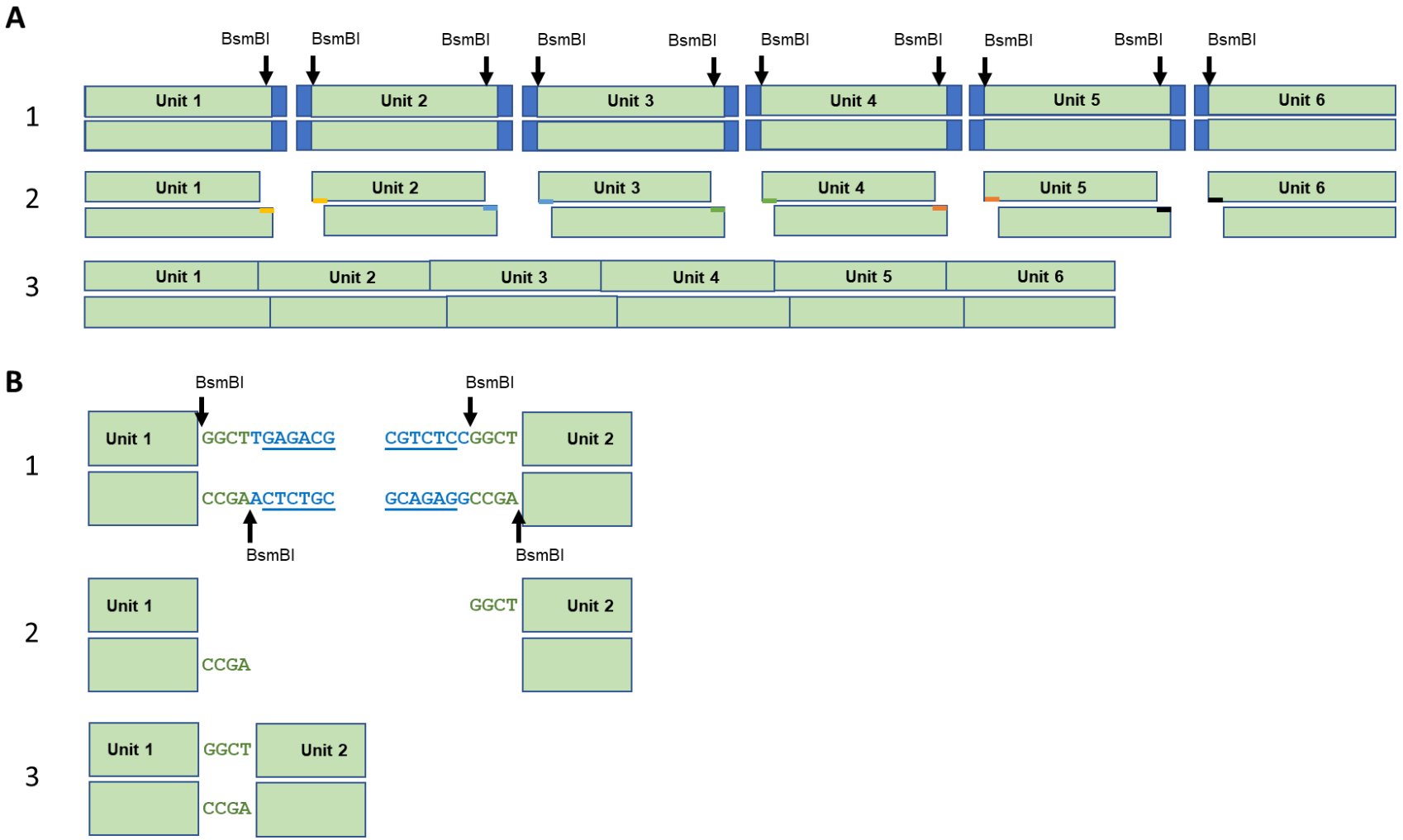

Figure 3.11. Schematic to show the ligation of units into a full-length construct. T1, constructed from units 1-6, is used as an example. 'A' illustrates an overview of the process, whereas 'B' focuses in more detail on the junction between units 1 and 2.
1: Units are digested with BsmBI: a type IIS restriction enzyme which recognises asymmetric DNA sequences and cleaves outside of its recognition sequence (in blue). The recognition sequence is thereby cleaved away from the unit, leaving only coding sequence (in green).
2: Restriction digestion reveals overhangs between units that are uniquely complementary (represented by yellow, blue, green, orange and black lines), therefore only allow units to assemble directionally and in the correct order.
3: BsmBI-digested units are incubated together in a T4 DNA ligase reaction, where the complementary overhangs anneal and the nicks are sealed. This results in a seamlessly ligated full-length construct.

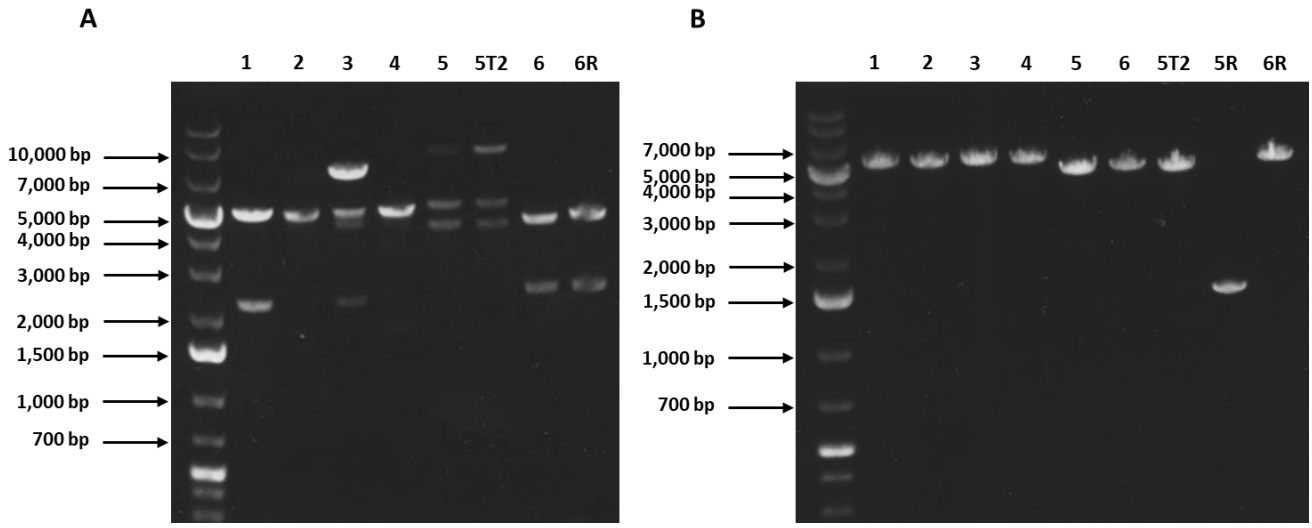

Figure 3.12. Analysis by agarose gel electrophoresis of units to be ligated into full-length constructs. All DNA was run on a 0.7% agarose gel at 100 V for 40 minutes and DNA bands were imaged under a Dark Reader transilluminator. The positions of relevant DNA mass markers are shown in bp. The number above each lane corresponds to the unit number (e.g. 1 is unit 1).
A: All units were digested with BsmBI to liberate them from their plasmids (or remove flanking sequences in the case of PCR products) and reveal uniquely compatible ends. Unit 5R is not shown in this image. Unit 3 appears to have digested partially.
B: The DNA bands corresponding to units were extracted from the gel, then an aliquot of each digested, extracted unit was run on another gel to check for size and purity.

To produce full-length constructs, the appropriate units were incubated together in a T4 DNA ligation reaction. The assembled constructs were analysed by agarose gel electrophoresis, which required optimisation due to the large size of the constructs (~29 kb for T1 and T2S and ~26 kb for the replicon cDNA). After trialling different conditions, it was found that running a TAE gel in TAE buffer at a low voltage (10-15 V) for 24 hours resulted in adequate resolution of large DNA bands. Under these conditions, agarose gel electrophoresis revealed a DNA band of the correct size for each construct in addition to many non-specific bands (Figure 3.13). In order to optimise the ligation reaction, different ligation temperatures and DNA amounts were trialled for the replicon construct, but there was no appreciable difference between conditions (Figure 3.14).

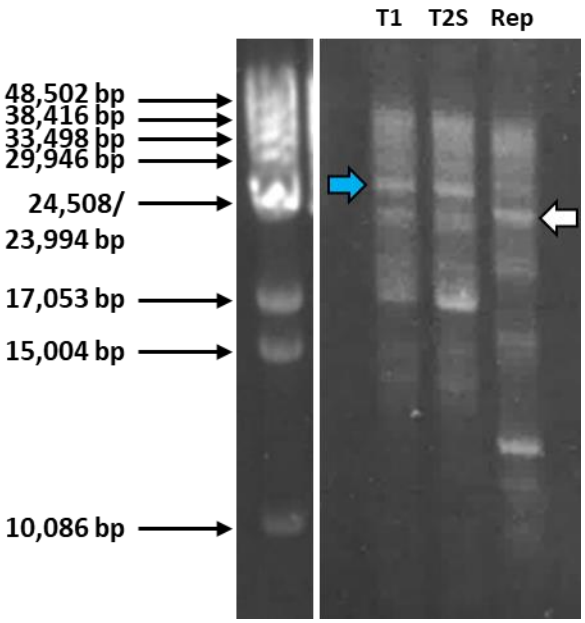

Figure 3.13. Analysis of the products of full-length construct ligation. Full-length constructs were ligated together from their constituent units in a T4 DNA ligation reaction. Ten percent of the ligation reaction was run on a 0.4% agarose gel at 15 V for 24 hours and imaged under a UV transilluminator to check for the presence of full-length product. Bands of approximately the correct size for full-length product (~29 kb for T1 and T2S, blue arrow; ~26 kb for replicon, white arrow) were present for all constructs, as well as many non-specific bands. The positions of relevant DNA mass markers are shown in bp. Rep: replicon.

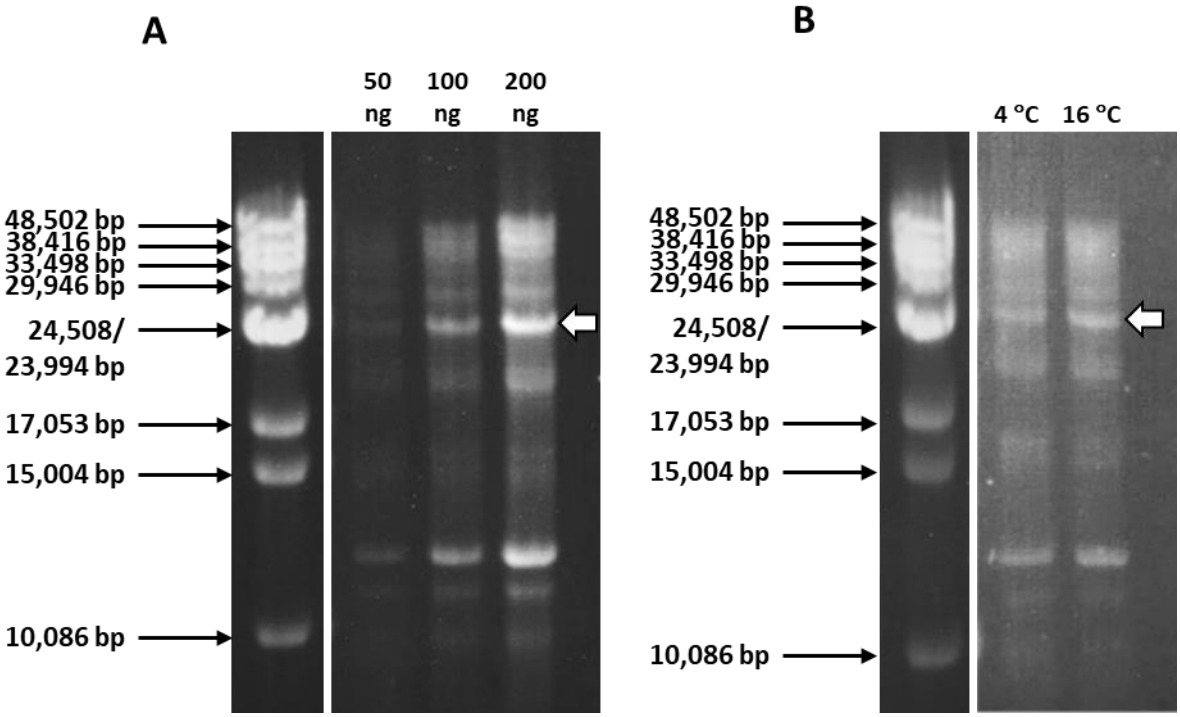

Figure 3.14. Comparison of different ligation conditions for replicon ligation. The replicon construct was ligated together from its constituent units in T4 DNA ligation reactions with different total DNA amounts (A) and ligation temperatures (B). Ten percent of the ligation reaction was run on a 0.4% agarose gel at 15 V for 24 hours and imaged under a UV transilluminator to compare the efficiency of the reaction between conditions. The band assumed to represent full-length replicon is indicated with a white arrow. The positions of relevant DNA mass markers are shown in bp.

In order to verify that the full-length constructs had been correctly assembled, a set of PCR primers was designed to produce seven overlapping amplicons from each assembled full-length construct, spanning the construct and ranging from 1.9 to 6.4 kbp in size (Figure 3.15 and Table 3.2). Amplicons of the correct size were achieved for all constructs except the replicon, from which 'amplicon 6' failed to be amplified. Additional primers were therefore designed to divide replicon amplicon 6 into two parts (6.1 and 6.2), which yielded amplicons of the correct size (Figure 3.16).

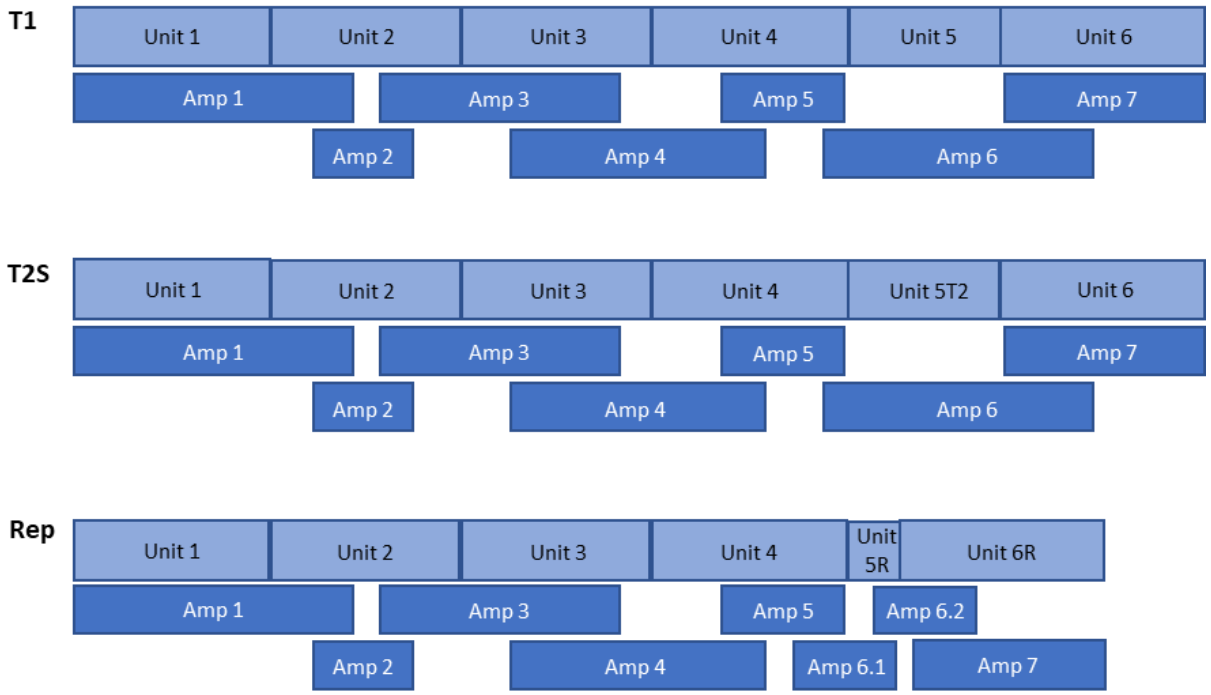

Figure 3.15. Schematic to show the amplicons obtained by PCR from the three full-length constructs. Constructs (T1, T2S and replicon (Rep)) are shown in pale blue, made up from their constituent units. The set of amplicons produced from each construct is shown beneath it in dark blue. The positions and sizes of the amplicons are approximately correct relative to the constructs. Amp: amplicon.

Table 3.2. The expected size of the PCR amplicons amplified from the three full-length constructs. Replicon amplicon 6 was divided into two parts: 6.1 and 6.2.

Amplicon	T1	T2S	Replicon
1	6,447 bp	6,447 bp	6,447 bp
2	1,949 bp	1,949 bp	1,949 bp
3	5,513 bp	5,513 bp	5,513 bp
4	5,902 bp	5,902 bp	5,902 bp
5	2,293 bp	2,293 bp	2,293 bp
6	6,105 bp	6,063 bp	1,563 bp (6.1)
			1,706 bp (6.2)
7	4,465 bp	4,465 bp	4,611 bp

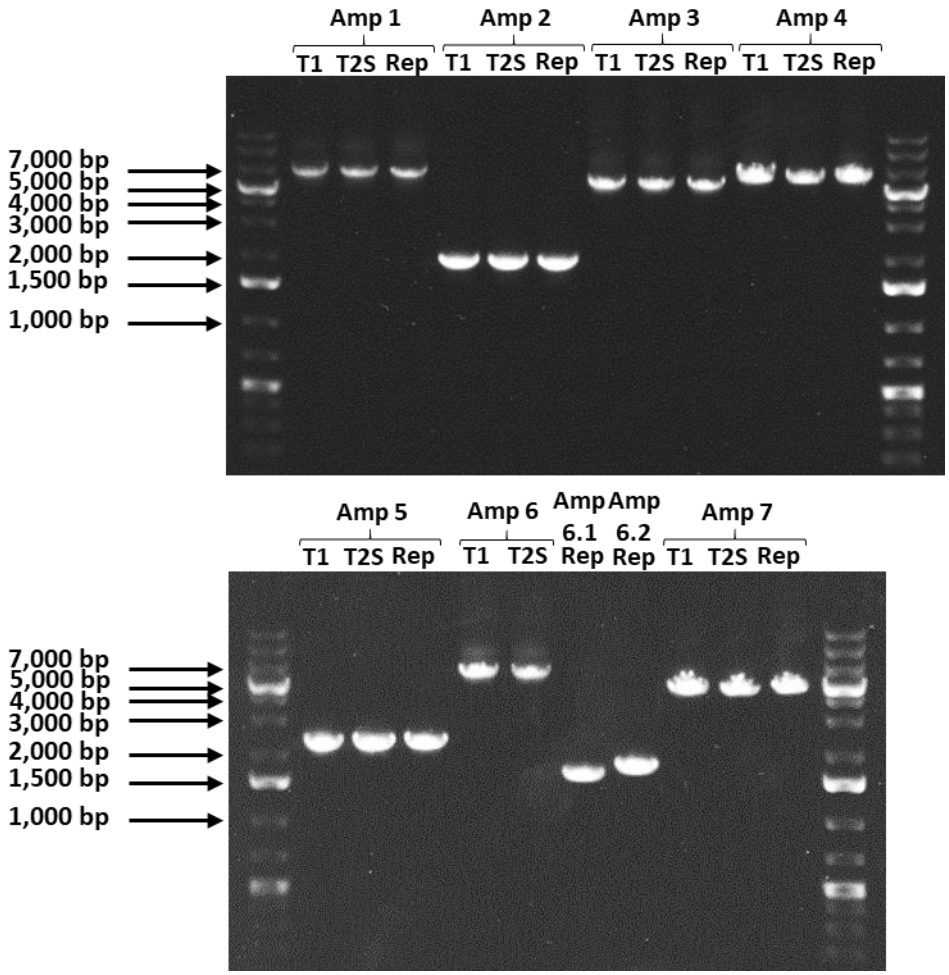

Figure 3.16. Analysis of the PCR amplicons obtained from each full-length construct by agarose gel electrophoresis. A set of PCR primers was designed to produce seven overlapping amplicons spanning each assembled full-length construct. Amplicons were run on a 0.7% agarose gel at 100 V for 40 minutes and imaged under a UV transilluminator. The number of the amplicon (Amp) and name of the construct from which it is derived is displayed above each lane. Replicon amplicon 6 was divided into two parts (6.1 and 6.2). The positions of relevant DNA mass markers are shown in bp. Rep: replicon.

3.2.7 Transcribing full-length constructs

The full-length constructs ligation reactions were purified by phenol chloroform extraction and ethanol precipitation and used as templates for *in vitro* transcription. An FCoV N gene DNA template was also created by PCR amplifying a ~1.2 kb region encompassing the N gene from unit 6, using primers that added a T7 promoter to the 5' end of the gene and a poly-A tail to the 3' end (Appendix A).

The first construct to be *in vitro* transcribed was replicon, which was transcribed alongside N gene and a positive control template included with the kit (pTRI-Xef). Analysis of the transcripts by spectrophotometer and agarose gel electrophoresis showed that no replicon RNA had been made, despite a reasonable yield of N gene RNA (~1,300 ng) and an excellent yield of positive control RNA (~30,000 ng) (Figure 3.17A). Although the same mass of replicon and N gene templates had been used in the reaction, the replicon template contained many DNA bands in addition to the full-length product so on a molar basis less full-length template would have been available for transcription.

The *in vitro* transcription reaction was therefore repeated using a higher concentration of replicon DNA template. Replicon RNA was transcribed this time (~800 ng) and a band was detected by agarose gel electrophoresis (Figure 3.17B). If the replicon RNA was ~30,000 nt as expected, the band should have migrated between the 10,000 and 20,000 bp DNA markers, but it was much larger than this. Since the agarose gel was not run under denaturing conditions, this could have been the result of secondary structure in the RNA. Replicon RNA was therefore examined by denaturing agarose gel electrophoresis (Figure 3.18), which showed a band of over 20,000 nt. It was, however, impossible to accurately size the full-length RNA as there is no RNA mass marker available that is large enough.

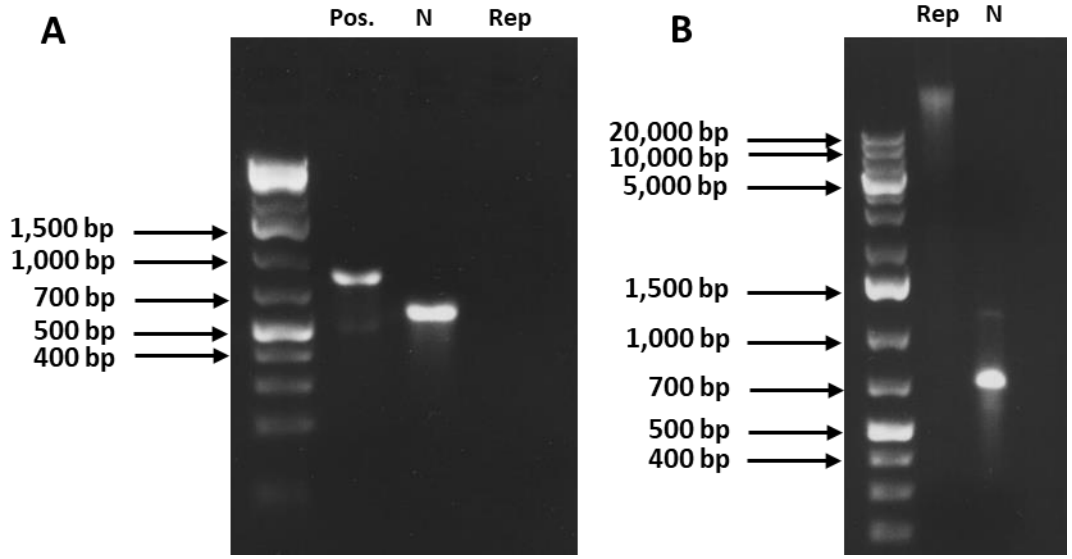

Figure 3.17. Analysis of replicon RNA by agarose gel electrophoresis. RNA, obtained by *in vitro* transcription of replicon (Rep), N gene (N) or positive control (Pos.) template DNA, was run on a 1% agarose gel at 100 V for 35 minutes. The RNA was denatured prior to loading but the gel was not run under denaturing conditions. Image 'A' shows RNA obtained from the first *in vitro* transcription reaction and image 'B' shows RNA obtained from the second. Note that a DNA (not RNA) ladder was used for sizing RNA, so RNA is approximately double the size that it appears to be. The positions of relevant DNA mass markers are shown in bp.

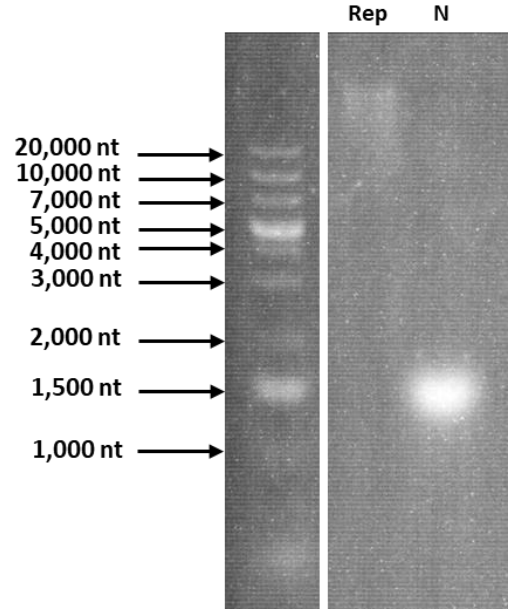

Figure 3.18. Analysis of replicon RNA by denaturing agarose gel electrophoresis. RNA, obtained by *in vitro* transcription of replicon (Rep) or N gene (N) template DNA, was run on a 1% agarose gel at 100 V for 35 minutes under denaturing conditions. The positions of relevant RNA mass markers are shown in nt.

RNA was transcribed from the T1 and T2S DNA templates under the same conditions used for the replicon template and the yields were ~700 ng and ~200 ng RNA respectively. Although RNA was being produced for the full-length constructs, each transcription reaction consumed a lot of DNA template and resulted in a yield that was lower than ideal for downstream applications. In an attempt to improve the efficiency of the reaction, *in vitro* transcription was carried out for four hours instead of two. However, this resulted in an even lower RNA yield. A strategy was therefore devised to increase the amount and purity of DNA template available for transcription, which involved designing primers complementary to the 5' and 3' ends of each construct (Appendix A), then using long-range PCR to amplify the full-length constructs from their ligation reactions. This yielded amplicons that appeared to be the right size by agarose gel electrophoresis for T1 and T2S (Figure 3.19), but no specific amplicon could be obtained for the replicon.

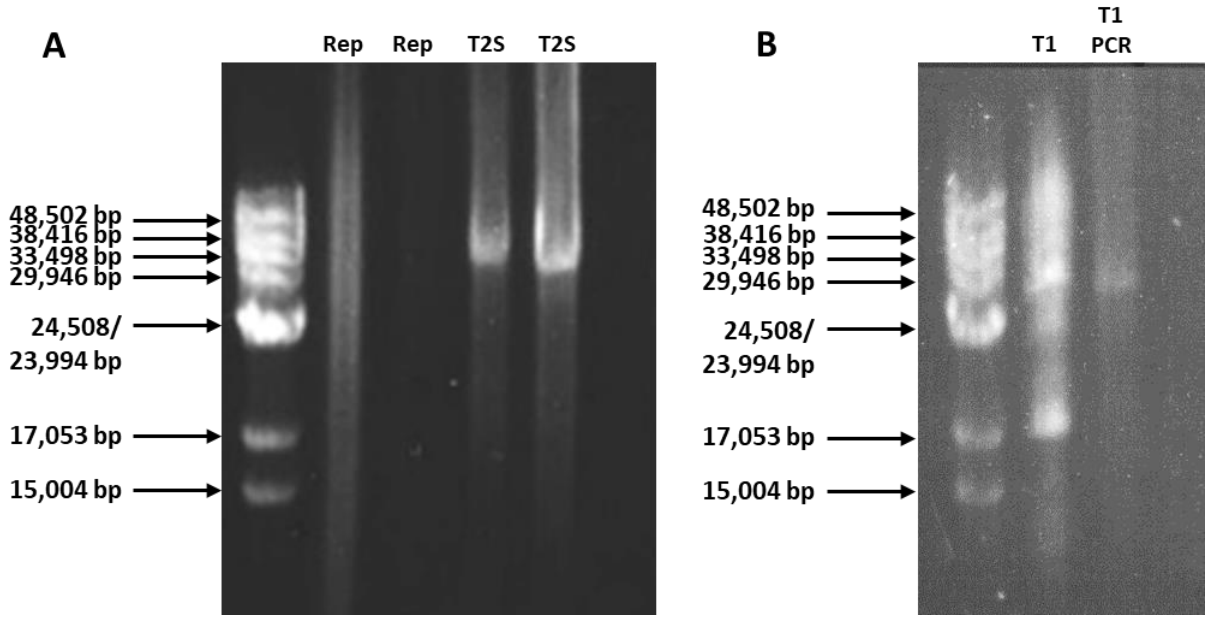

Figure 3.19. Analysis of long-range PCR amplicons by agarose gel electrophoresis. Primers were designed to amplify each full-length construct from its respective ligation reaction using long-range PCR. Amplicons were run on a 0.4% agarose gel at 15 V for 24 hours and imaged under a UV transilluminator. The name of the construct from which the amplicon is derived is displayed above each lane. The positions of relevant DNA mass markers are shown in bp.
A: Long-range PCR was carried out to amplify replicon (Rep) and T2S in duplicate.
B: Long-range PCR was carried out to amplify T1 (T1 PCR), and this was run alongside T1 full-length construct DNA (i.e. not PCR-amplified) for comparison.

T1 and T2S long-range PCR products were used as templates for transcription and resulted in an approximately 40-fold increase in RNA yield, but when analysed by denaturing agarose gel electrophoresis the RNA transcribed from the PCR-amplified templates was smaller than that transcribed directly from the ligated template (Figure 3.20). Although an RNA marker large enough to verify the size of the full-length RNA was not available, it was assumed that RNA made from PCR-amplified template DNA was truncated. This was confirmed when RT-PCR failed to amplify the 3' end of the RNA (data not shown). Transcription from this point was only carried out using template that had not been PCR-amplified, leading to a lower yield of RNA but a presumably full-length product.

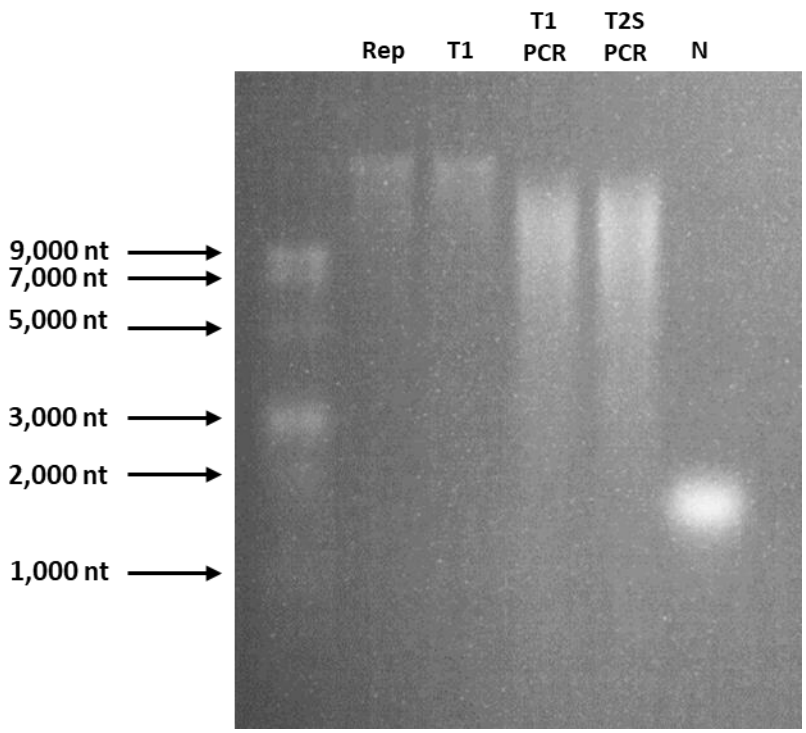

Figure 3.20. Analysis of full-length construct RNA by denaturing agarose gel electrophoresis. RNA, obtained by *in vitro* transcription of full-length construct template DNA, full-length construct template DNA amplified by long-range PCR or N gene template DNA, was run on a 1% agarose gel at 100 V for 40 minutes under denaturing conditions. The name of the construct is displayed above each lane. The positions of relevant RNA mass markers are shown in nt. Rep: replicon, T1 PCR: T1 transcribed from PCR-amplified template, T2S PCR: T2S transcribed from PCR-amplified template, N: N gene.

3.2.8 Transfecting cells with full-length construct RNA

With each construct transcribed, work began to transfect the RNA into mammalian cells and produce infectious virus or replicon. In all cases, cells were co-transfected with full-length construct and N gene RNA, because inclusion of N gene transcripts has been shown to promote the infectivity of coronavirus infectious clones assembled by *in vitro* ligation (Yount *et al.*, 2000).

3.2.8.1 Transfecting cells with T2S RNA

T2S RNA was transfected into BHK cells under three conditions: T2S RNA only, T2S+N gene RNA at a ratio of 1:1, and N gene RNA only. BHK cells were chosen because they lack an IFN response (Otsuki *et al.*, 1979) so may be less resistant to foreign RNA. However, BHK is a hamster cell line that does not express feline APN, so even if infectious T2S virions were produced, the infection would be abortive in BHK cell culture. The culture supernatant from cells transfected with T2S+N gene was therefore passaged on to CrFK cells (known to express feline APN) for two passages. The BHK cells that were originally transfected showed no cytopathic effect. On examination by immunofluorescence assay using an anti-coronavirus antibody (FIPV3-70) raised against CCoV but known to have reactivity with FCoV N protein (Poncelet *et al.*, 2008), no transfected BHK cells were stained by the antibody (Figure 3.21), suggesting that coronavirus antigen was not present. A positive control for the assay was not available, so it could not be ruled out that infectious T2S had been produced but was not being detected. However, FIPV3-70 should have been capable of recognising N protein had it been translated from the N gene RNA. To further explore whether the transfection was successful, RNA was extracted from an aliquot of cell culture supernatant taken at each passage. The RNA was reverse transcribed and a PCR was carried out on the resulting cDNA using primers (Appendix A) to generate a 2.2 kb amplicon from the FCoV ORF1a-b gene. The PCR products were analysed by agarose gel electrophoresis, but no DNA band was seen for any passage (Figure 3.21). These results suggest that T2S RNA did not yield infectious virus.

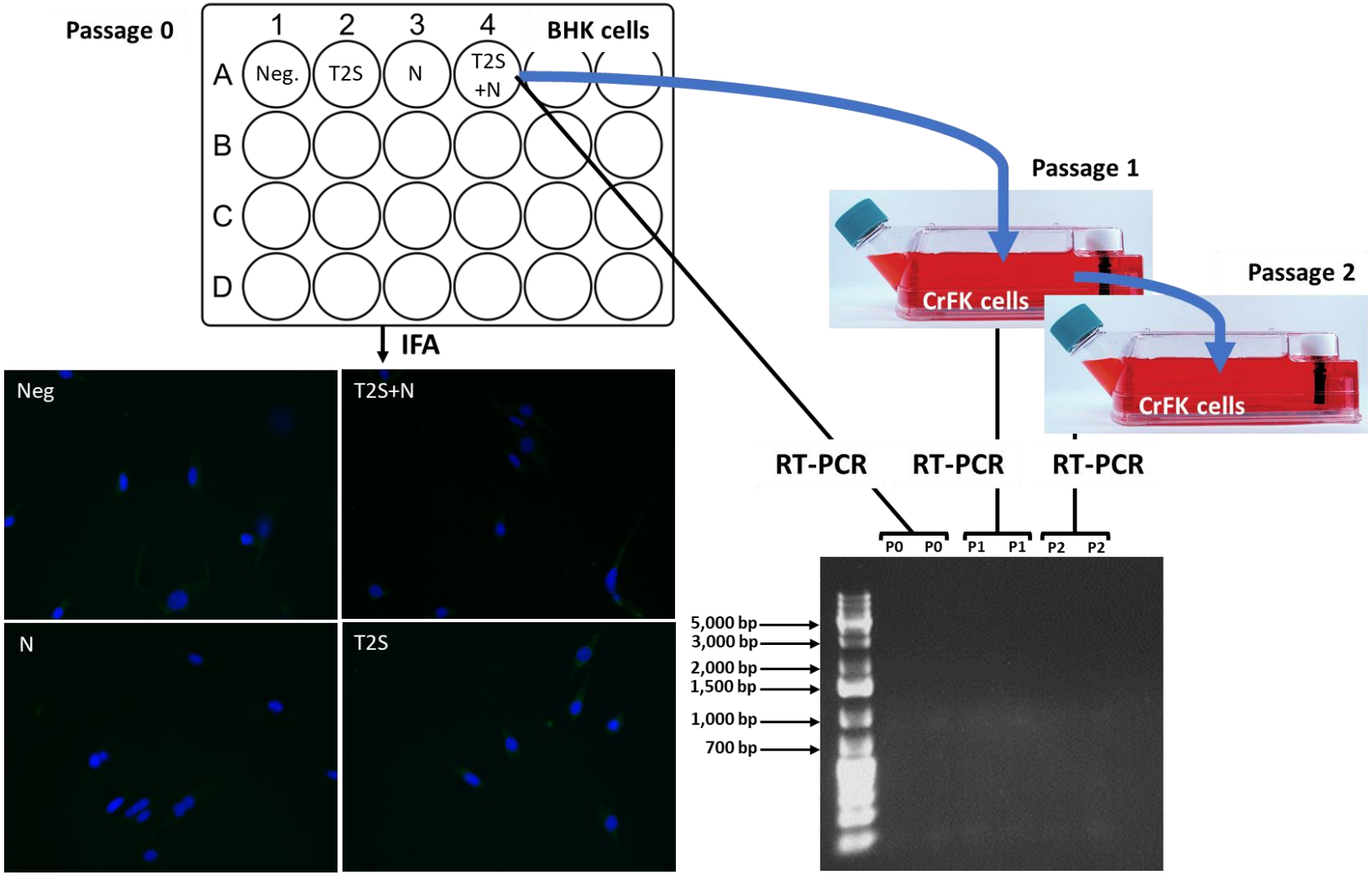

Figure 3.21. Schematic to illustrate the T2S transfection strategy, plus the results of the transfection. BHK cells were transfected with T2S, N gene (N), T2S and N gene (T2S+N) or no (Neg.) RNA (Passage 0). The supernatant of the cells transfected with T2S+N gene RNA was passaged on to CrFK cells (Passage 1). The supernatant of the Passage 1 CrFK cells was passaged on to more CrFK cells (Passage 2). The cells that were originally transfected underwent an immunofluorescence assay (IFA), in which they were fixed with permeabilisation and incubated with an anti-coronavirus antibody FIPV3-70, followed by a DyLight 488-conjugated secondary antibody (green). Nuclei were stained with DAPI (blue). Cells were then viewed using a widefield fluorescence microscope with a 40x oil immersion objective. An aliquot of cell culture supernatant was taken at each passage (P0-2) and the RNA was extracted. An RT-PCR was carried out on the RNA, in duplicate, using primers that would yield a 2.2 kb amplicon if FCoV was present. A positive control for the PCR was not included. PCR products were run on a 0.7% agarose gel at 100 V for 40 minutes and imaged under a UV transilluminator.

3.2.8.2 Transfecting cells with replicon RNA

In an attempt to establish a replicon cell line, cells were transfected with replicon RNA under three conditions: replicon+N gene RNA at a ratio of 4:1, replicon+N gene RNA at a ratio of 1:1, and N gene RNA only. The transfection was first carried out using CrFK cells, since they are known to support growth of Type 2 FCoV (Van Hamme *et al.*, 2007) so should allow replication of the viral genome. Following transfection, puromycin was applied at a concentration of 5 μg/ml to select for cells carrying the replicon (which encodes puromycin resistance). This puromycin concentration was selected based on a survey of the literature and an experiment demonstrating that this was the lowest concentration of puromycin that killed 100% of untransfected CrFK cells (data not shown). No cells survived following puromycin selection and no expression of GFP (also encoded by the replicon) was observed in dying cells. This result suggests that the replicon RNA had not undergone replication to produce sgRNA, or if produced the sgRNA was not translated to detectable levels.

The transfection was repeated twice using BHK cells. Following the first transfection of BHK cells, 3 μg/ml puromycin was applied to select for cells carrying the replicon. This puromycin concentration was selected based on the protocols previously used in the laboratory. Since no cells survived following puromycin selection, a second transfection was carried out and 1.5 μg/ml puromycin was used for selection. This was to avoid the possibility that the replicon was being expressed, but the puromycin concentration was too high for the puromycin resistance protein to overcome it. Again, no cells survived following puromycin selection and no expression of GFP was observed in dying cells, suggesting that the replicon RNA either had not undergone replication to produce sgRNA or they were not translated in BHK cells.

3.2.9 Next generation sequencing

Since transfection with the full-length constructs failed to produce infectious virus or replicon-expressing cells, next generation sequencing was used to verify the sequence of the RNA transcripts. Full-length *in vitro* RNA transcripts were reverse transcribed and PCR amplicons were generated from the resulting cDNA using the PCR primers described in section 3.2.6 (Figure 3.22). Amplicons of the correct size were achieved for all constructs.

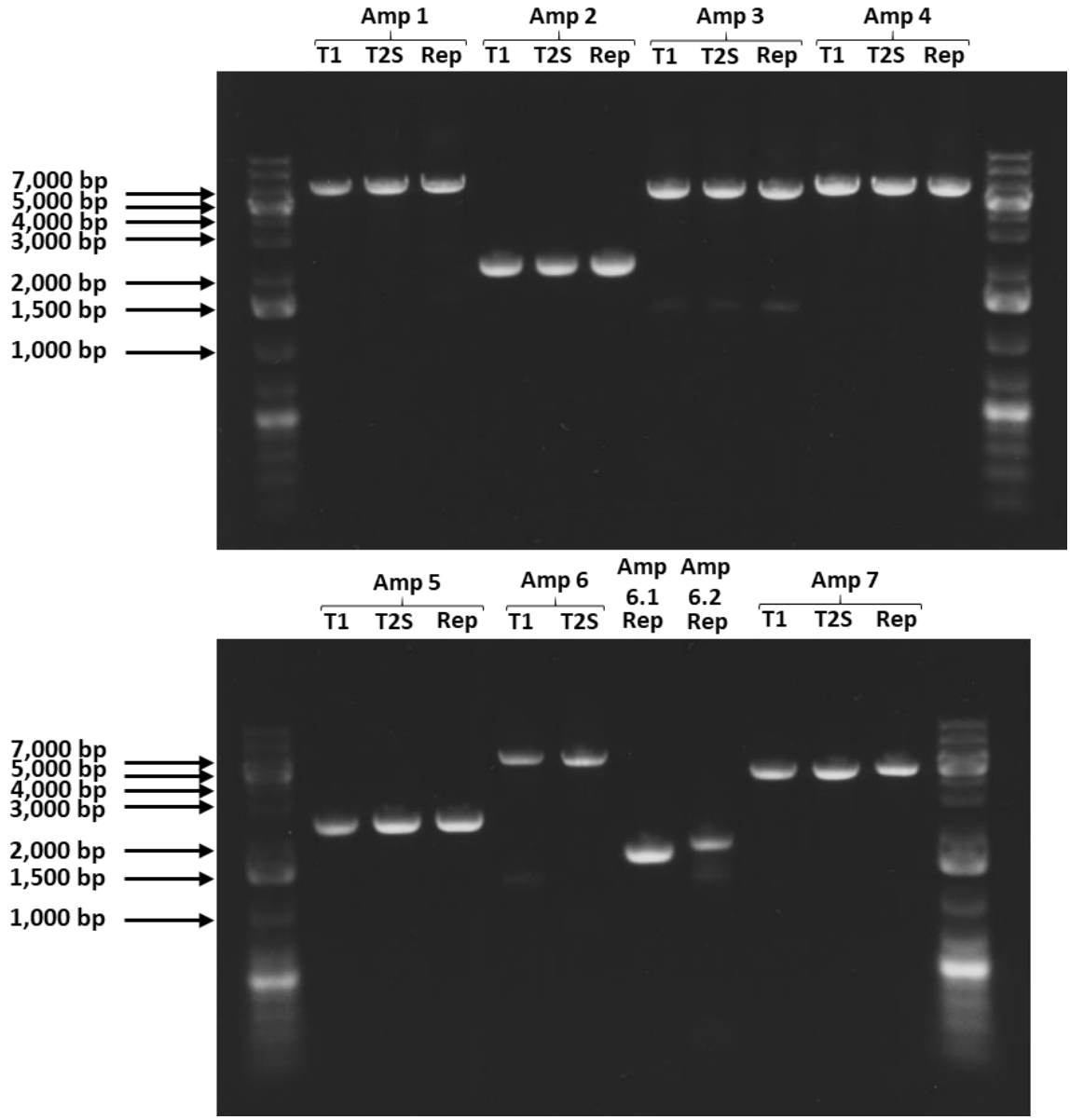

Figure 3.22. Analysis of the RT-PCR amplicons obtained from each full-length construct RNA by agarose gel electrophoresis. A set of PCR primers was designed to produce seven or eight overlapping amplicons spanning each full-length *in vitro* RNA transcript, following a reverse transcription step. Amplicons were run on a 0.7% agarose gel at 100 V for 40 minutes and imaged under a UV transilluminator. The number of the amplicon (Amp) is displayed above each lane. Replicon (Rep) amplicon 6 was divided into two parts (6.1 and 6.2). The positions of relevant DNA mass markers are shown in bp.

cDNA amplicon pools were created for each construct and these were submitted to the Genomics Facility, University of Bristol, for library preparation and next generation sequencing (Figure 3.23). The sequencing results showed that there was full coverage of each construct with a minimum

depth of 2,625 reads at each base, though there were spikes in read depth corresponding to PCR primer annealing sites. The consensus genome of the replicon construct was correct, but the T1 and T2S constructs contained three and two areas respectively of sequence that deviated from expected. Moving from 5′ to 3′, the first area of sequence deviation, common to both T1 and T2S, consisted of one base and lay within amplicon 1, derived from unit 1. The second area, also common to both T1 and T2S, consisted of six bases and lay within amplicon 3, derived from unit 2. The third area, only present in T1, consisted of six bases and lay within amplicon 6, derived from unit 5. None of the areas overlapped with PCR primer annealing sites or unit boundaries. In each of these areas there was heterogeneity in the sequence, with a significant proportion (in most cases the majority) of reads representing the correct sequence and the remainder representing the incorrect sequence. It is unclear what these areas of heterogeneity represent, but the results suggest that the constructs were generally ligated and transcribed correctly to produce RNA faithful to the expected sequence. It is therefore likely that the attempts to produce infectious virus or replicon-expressing cells following transfection of the RNA into cells failed due to a different reason.

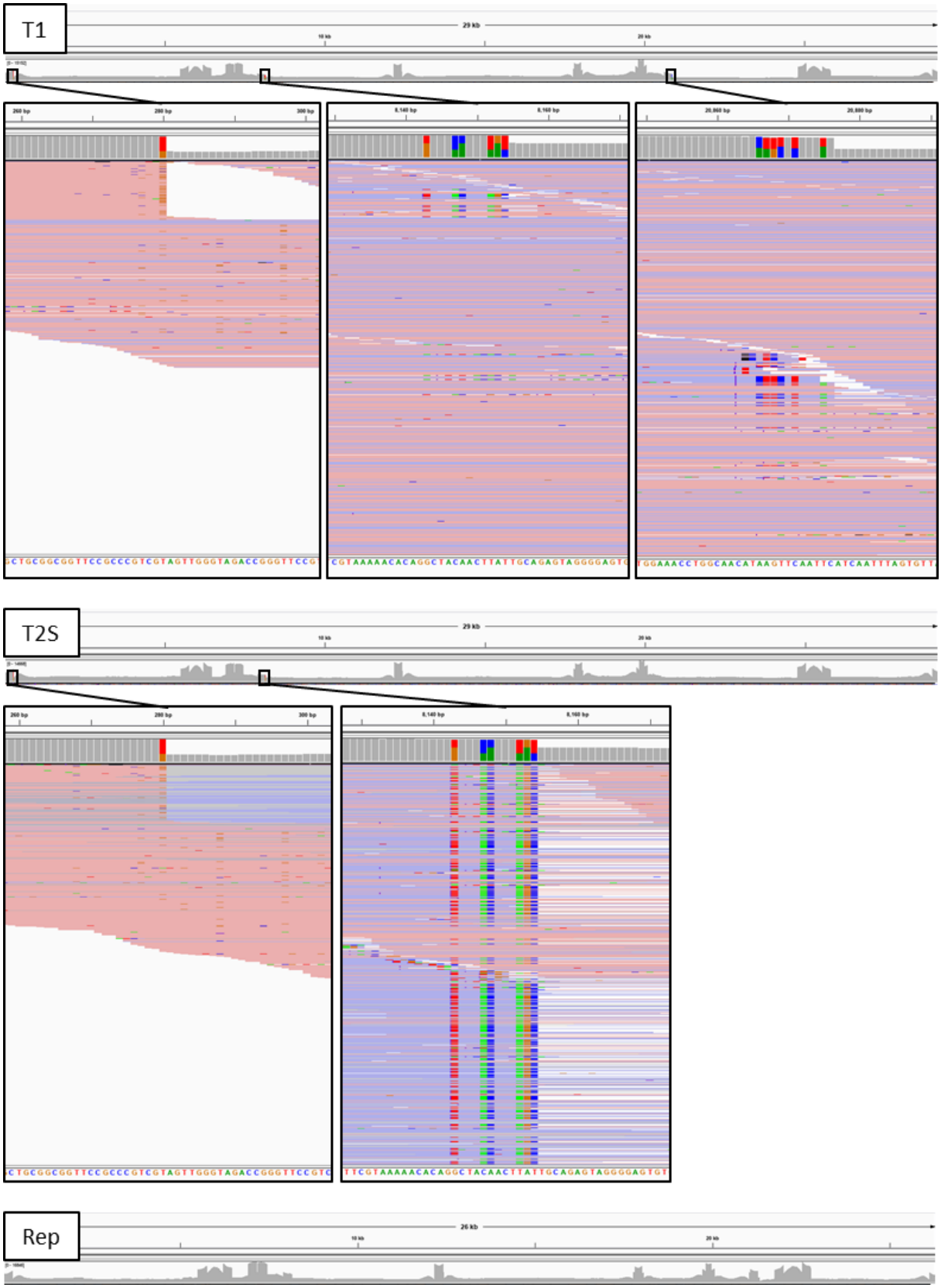

Figure 3.23. Next generation sequencing results of the three full-length constructs. Full-length RNA transcripts of T1, T2S and replicon (Rep) were reverse transcribed and PCR amplified to produce a set of cDNA amplicons ranging from 1.5 to 6.5 kb. The set of amplicons representing each construct was pooled in an equimolar fashion and submitted for library preparation and next generation sequencing. The reads were aligned against the expected sequence of each construct. Read depth is shown as a grey trace. Where there is significant deviation from the expected sequence, that base position is shown in colour instead of grey. A closer view of these areas reveals the proportion of reads at the position representing the deviated sequence, with the colour corresponding to the base (red: T, blue: C, green: A, brown: G). The expected sequence is displayed at the bottom.

3.3 Discussion

This chapter describes the steps taken towards establishing a Type 1 FCoV reverse genetic system. Such a system would allow mutations to be strategically introduced into the viral genome, enabling investigation of the consequences of given mutations on virus phenotype. It would also serve as a powerful platform for vaccine development and antiviral drug screening.

An *in vitro* ligation approach was taken to produce the full-length infectious cDNA clones in this project. This method was adopted because it was potentially fast and simple compared to the alternatives of the use of vaccinia virus or BAC vectors. Additionally, this method has been used successfully to recover a number of other recombinant coronaviruses (Almazan *et al.*, 2014). The main obstacle to the production of the full-length constructs built was the instability of certain regions of the genome when undergoing amplification in bacteria, leading to deletions, mutations and rearrangements. This was particularly the case with units 1, 2 and 4, which all correspond to regions within the FCoV ORF1a-b gene. This region has previously been reported to contain sequences that are unstable in bacterial culture (Scobey *et al.*, 2013, Yount *et al.*, 2000), including in BAC vectors (Thiel *et al.*, 2001). Considering that methods that use vaccinia virus as a vector also require some degree of *in vitro* amplification of genome fragments (Almazan *et al.*, 2014), it is likely that this problem would have arisen regardless of the method used to produce the full-length infectious cDNA clones.

To overcome the problem of instability, two strategies were utilised: cloning into a low copy number vector and PCR amplification. The only units that were successfully cloned into low copy number vectors were units 5 and 5T2, and this was out of necessity because the original strategy of cloning all fragments into the fragment 'A' plasmid was not possible. A low copy number vector, pWSK29, was chosen because it was assumed – given that their 'A' fragments could not be amplified successfully in bacteria – that units 5 and 5T2 might contain sequences toxic to bacteria. This vector in particular was chosen because it has been used to overcome instability issues when assembling full-length infectious cDNA clones of SARS-CoV (van den Worm *et al.*, 2012). Primers were designed to add flanking sequences on to units 2 and 4, but these units were never cloned into the low copy number vector and instead were amplified by PCR before being ligated into the full-length constructs. Given more time, it would be beneficial to clone these units into the low copy number vector so they are available as stable plasmids that can be amplified in bacterial culture.

The approach taken in this project was to clone fragments and then units into plasmids and amplify them in bacterial culture, only using PCR amplification as a last resort. This is because PCR results in DNA polymerase-induced errors in the amplified DNA sequence. However, the error rate is

extremely low with proofreading polymerases (Cha and Thilly, 1993). Given the problems experienced over the course of this project with amplifying DNA in bacterial culture, using PCR to amplify from the beginning would have made the process of assembling full-length infectious clones easier. As the project progressed, PCR was increasingly used successfully to amplify fragments (1B) and units (2 and 4) that were unstable in bacterial culture. In future, PCR should be considered as a strategy for amplifying unstable fragments and units from the beginning.

Eventually all units taken forward to be ligated into full-length constructs were sequence-perfect except for unit 1, where three mutations were introduced when the ligated unit was grown in bacterial culture. The mutations were T1796C (T>C at unit 1 position 1796), T2225C (T>C at unit 1 position 2225) and A2487G (A>G at unit 1 position 2487). The first two mutations are synonymous so would not be expected to influence the structure or function of the final infectious clone into which the unit would be incorporated, but the latter mutation results in an amino acid substitution of valine for isoleucine within the ORF1a gene. Both are aliphatic amino acids with no charge, but valine is larger and more hydrophobic than isoleucine. In order to assess the conservation of isoleucine at the position of the substitution, the 70 most similar sequences to FCoV 80F, comprising strains of FCoV and CCoV, and two TGEV strains were analysed. All 71 sequences were found to have an isoleucine at the position of the substitution, suggesting that this amino acid is highly conserved and therefore may be critical for coronavirus replication. However, the position of the substitution suggests that, once ORF1a is translated and cleaved, it would sit within nsp2, which is thought to interfere with host defences rather than support viral replication directly (Neuman *et al.*, 2014). It is unknown whether this amino acid substitution will impact the infectious clones, but, given more time, it would be preferable to repair the mutation. A possible strategy for doing this would be to PCR amplify the unit from the NEBuilder HiFi DNA assembly ligation reaction, as was done successfully for unit 4. This should result in a sequence-perfect unit 1, as the bacterial culture step that presumably introduced the mutations would be avoided.

The full-length constructs were designed to encompass flanking sequences that would allow them to be ligated into a vaccinia virus vector. Given that one of the limitations experienced in the project was generating enough full-length construct DNA template to get a good yield of RNA, transferring the constructs into vaccinia virus seems like an obvious next step to overcome this challenge. Within this project, attempts were made to optimise the full-length construct ligation step by changing the DNA concentration and temperature of ligation reactions, neither of which made an appreciable difference; as well as a strong band corresponding to full-length DNA template seen on agarose gel electrophoresis of the ligation reactions, many non-specific bands were present. These bands likely

represented units that had joined together incorrectly, though this should not have been able to happen since the BsmBI sites between the units were unique.

Long range PCR was briefly tested as a way of amplifying the full-length DNA template and thereby increasing RNA yield. At first this appeared to work extremely well for the T1 and T2S constructs, with a 40-fold increase in RNA yield when transcribing PCR-amplified template compared with non-amplified. However, it soon emerged that the RNA transcribed from PCR-amplified template was truncated at the 3′ end. The truncation was discovered when PCR failed to produce an amplicon from the 3′ end of PCR-amplified T2S and T1. Unfortunately, this only occurred after an attempt was made to transfect cells with T2S RNA made from PCR-amplified template. It is therefore not surprising that the transfection failed to yield infectious virus. In future work this experiment should be repeated using T2S RNA transcribed from a non-amplified template, ideally including a positive control. In this experiment, transfection success was measured in two ways: an immunofluorescence assay examining the transfected cells for coronavirus antigen, and RT-PCR examining the cell culture supernatant at each passage for FCoV RNA. A positive control for both assays could be RNA of a different coronavirus infectious clone that is known to yield infectious virus in BHK cells. A positive control for the PCR element of the RT-PCR could be any of the full-length DNA templates produced in this project.

Attempts were also made to transfect cells with replicon RNA. Despite the RNA being transcribed from non-amplified template, the experiment failed in two different cell types (CrFK and BHK). CrFK cells were used first because they are feline cells known to support propagation of Type 2 FCoV (Van Hamme *et al.*, 2007), but no cells expressed GFP or survived following puromycin selection. It was decided to use BHK cells instead, because they lack an interferon response (Otsuki *et al.*, 1979) and therefore may be more permissive to the replicon. The transfection was repeated twice using BHK cells, the second time using a very low puromycin concentration to ensure that cells expressing the replicon were not being killed. However, no cells expressed GFP or survived following puromycin selection on both occasions. As a positive control for transfection was not included, it is unknown whether the experiments failed at the transfection stage or later. In the future a positive control mRNA (e.g. encoding a reporter gene) could be included in the transfection protocol. Almazan *et al.* (2004) demonstrated that the same replicon construct can be cytopathic to some cell types and non-cytopathic to others. It is possible that the replicon was being expressed but was cytopathic in the cell lines used, although, if this was the case, GFP expression should have been observed in cells before they died.

Cells were co-transfected with full-length construct and N gene RNA because inclusion of N gene transcripts has been shown to improve recovery rates and replication efficiency of infectious clones (Beall *et al.*, 2016, Yount *et al.*, 2000). The anti-coronavirus antibody (FIPV3-70) used in this project is known to specifically recognise the FCoV N protein (Poncelet *et al.*, 2008). If the N gene transcripts had been translated, the cells transfected with N gene (with or without T2S) should have been positive with this antibody in an immunofluorescence assay. However, this was not the case. The N gene is a simple construct that was available in this project at a high purity and concentration, so its expression in transfected cells should have been straightforward. The failure to detect N protein in an immunofluorescence assay may therefore indicate an issue with the transfection procedure, although this cannot be confirmed since a positive control for the assay was not included.

In this project a chemical method of transfection which had previously been used successfully for other coronavirus RNA transcripts was used. If further investigations suggested that the experiment was failing at the transfection stage, it would be worthwhile to attempt electroporation of *in vitro* RNA templates into cells instead. Electroporation is known generally to be more efficient than chemical methods for the introduction of genetic material into cells (Sharifi Tabar *et al.*, 2015), and electroporation has been used successfully for transfection of coronavirus full-length constructs into cells by other groups (Beall *et al.*, 2016, Ehmann *et al.*, 2018).

Given that neither infectious virus nor replicon-expressing cells were produced in the transfection experiments, it was decided to sequence cDNA fragments derived from the full-length *in vitro* RNA transcripts to check whether the constructs had been ligated and transcribed properly. A next-generation sequencing approach was taken because this was how Type 1 FCoV 80F was sequenced when it was first isolated from a clinical sample (Lewis *et al.*, 2015), therefore it is an established method and primer sequences were available for generating the amplicons. Many of the original primer sequences were used in this project, though some were changed to enable the same set of primers to produce amplicons from all three constructs. The primers annealing to the 5' and 3' extremes of the constructs were also redesigned to incorporate the entire sequence. Following reverse transcription, a full set of cDNA amplicons of the correct size was obtained for all constructs with minimal optimisation, suggesting that they had been ligated and transcribed properly. The amplicons were pooled for each construct and submitted for next-generation sequencing, which found that the constructs were generally as expected, corroborating that they had been ligated and transcribed correctly.

It is unclear what caused the areas of sequence heterogeneity observed in the T1 and T2S constructs, but, considering how many steps there were between ligating the sequenced units (i.e.

the last point where the sequence was verified) and next-generation sequencing, it is perhaps not surprising that errors were introduced. Due to the nature of the errors, it seems most likely that RNA secondary structure was interfering with the ability of the polymerase (either RNA polymerase in the transcription step or DNA polymerase in the reverse transcription step) to incorporate bases correctly. However, it is impossible to pinpoint at exactly which step the errors arose, therefore it is unknown whether some copies of the full-length construct RNA genuinely contained incorrect sequence or whether this was an artefact of the RT-PCR, library preparation or next-generation sequencing steps.

Nanopore sequencing was considered as an alternative to next-generation sequencing, since it is available and widely used in our laboratory. The advantages of nanopore sequencing are that direct RNA sequencing is possible and there is the capability of producing full-length reads. These features would avoid the need for reverse transcription of RNA, PCR of cDNA and fragmentation of amplicons, saving time and effort as well as circumventing steps that may have introduced errors. Additionally, this method can detect diverse and novel sgRNAs and methylation sites (Viehweger *et al.*, 2019). However, the main drawback of nanopore sequencing is a high base call error rate, which can be overcome when next-generation sequencing data are also available (Laver *et al.*, 2015). It could therefore be useful to carry out nanopore sequencing on the full-length construct RNA to verify the results of next-generation sequencing. Looking to the future, nanopore sequencing could also prove a useful tool to characterise FCoV transcripts produced in infected cells.

At the start of this project, the only published Type 1 FCoV reverse genetic system was based on the FCoV strain 'Black', which is laboratory-adapted and cannot be said to represent field strains of Type 1 FCoV (Pedersen, 2009). The need for a reverse genetic system based on a field strain of Type 1 FCoV was obvious, hence the decision to base our reverse genetic system on Type 1 FCoV 80F: a strain found in the faeces of a naturally-infected, healthy cat (Lewis *et al.*, 2015). Ehmann *et al.* (2018) since developed the first full-length infectious clone based on a field strain of Type 1 FCoV (recFECV). The group cloned a cDNA copy of the virus, originally isolated from the faeces of a healthy FCoV-infected cat, into a vaccinia virus vector. Like the T2S construct produced in this project, they also produced a chimeric construct with a Type 1 FCoV field strain backbone but a Type 2 FCoV '79-1146' S gene. Since recombinant FCoV carrying the '79-1146' S gene is known to infect FCWF cells (Tekes *et al.*, 2012, Thiel *et al.*, 2014), the chimeric construct was used in *in vitro* assays as a proxy for recFECV. The group found that infectious recFECV virions could be recovered from cell culture supernatant and, although these particles were not infectious *in vitro*, they caused productive infections in cats that closely mirrored what was seen with the parental virus. Though very similar, it

is still worthwhile to persevere with the reverse genetic system developed in this project because it offers two novel points:

1. FCoV mutates readily, so two strains with the same pathotype (i.e. FECV) can be genetically distinct, particularly when isolated from different geographical regions and at different time points (the Type 1 FCoV 80F used in this project was isolated in the UK in 2011 (Lewis *et al.*, 2015), whereas the strain used by Ehmann *et al.* (2018) was isolated in Germany in 2012). The production of infectious clones based on a variety of Type 1 FCoV field strains will lead to a fuller understanding of the virus.
2. This project aims to establish the first FCoV replicon.

A major limitation with the study by Ehmann et al. (2018), and something that has long impeded FCoV research, is absence of a cell line in which Type 1 FCoV can be propagated. Due to this, the authors had no choice but to employ *in vivo* propagation by using recFECV to directly infect live cats, when an *in vitro* characterisation stage may have been informative. It was hoped that, by identifying a Type 1 FCoV CER and producing a cell line capable of propagating Type 1 FCoV (see Chapter 4), this project would not only establish a Type 1 FCoV reverse genetic system but also a way of rescuing recombinant viruses *in vitro*. Part of the reason for developing T2S and replicon constructs in this project is that they would be usable even before a cell line supporting Type 1 FCoV propagation is developed. As well as helping to assess the success of the reverse genetic system, T2S would be useful for exploring the role of the S protein and investigating mutations in parts of the genome outside of the S gene. Similarly, the replicon would be a useful tool for investigating the replicative machinery of the virus and the effect of viral replication on the transcriptome and proteome of the cell, and would provide a fast, high-throughput way of screening anti-viral compounds.

4 Identifying a CER for Type 1 FCoV

4.1 Introduction

This chapter will describe the validation, optimisation and application of a method of identifying a Type 1 FCoV CER.

There are two FCoV serotypes: Type 1 and Type 2. Though Type 1 FCoV is responsible for most natural infections, Type 2 is the predominant serotype used for *in vitro* work. This is because there are many cultured cell lines available in which Type 2 FCoV can be propagated, whereas there are no known available cell lines in which Type 1 FCoV can be grown *in vitro*. On identification of a CER for Type 1 FCoV, the receptor could be stably expressed in a cell line, potentially rendering the cells permissive to infection. Such a tool would allow strains of FCoV representative of natural infections to be propagated *in vitro*, thus leading to a better understanding of the biology of the virus. A cell line which could be used to propagate Type 1 FCoV would also be a crucial component of a reverse genetic system, as it would allow rescue of recombinant viruses. A Type 1 FCoV reverse genetic system would enable the effects of viral mutations on phenotypic characteristics including tropism and virulence to be ascertained, and eventually serve as a platform for vaccine development and drug discovery.

The method used for identifying a Type 1 FCoV receptor in this project will be referred to herein as the 'bait protein' method. Raj *et al.* (2013) first used the bait protein method to identify a CER for the emerging MERS-CoV. The method involves expression of a chimeric protein comprising the S1 domain of the viral S protein and the Fc region of human IgG (i.e. the bait protein) and incubating the bait protein with cells susceptible to infection. The bait protein can be purified using its IgG Fc portion, and the cell surface protein to which the S1 portion binds can then be identified using mass spectrometry. Previously in the laboratory, bait proteins bearing the S1 domains of Type 1 (Type 1 bait protein) and Type 2 (Type 2 bait protein) FCoVs were produced and preliminary experiments done to optimise the expression and purification of the bait proteins for use in receptor identification experiments (Quirke, 2016). The S1 domain of the Type 1 bait protein is based on the FCoV strain 80F, which was isolated from the faeces of a naturally infected cat without FIP (Lewis *et al.*, 2015). The S1 domain of Type 2 bait protein is based on FCoV 79-1146, a highly virulent laboratory-adapted strain (Pedersen, 2009). In this project, the Type 2 bait protein will be used with CrFK cells – known to express the Type 2 FCoV CER, fAPN – to optimise and validate the bait protein method.

In the case of Raj *et al.* (2013), an approximately 110 kDa protein was precipitated from MERS-CoV susceptible Huh7 (human liver) and Vero (African green monkey kidney) cells by a bait protein bearing the S1 of MERS-CoV. This protein was visualised as a band on a gel, then the band was excised and identified as DPP4 by mass spectrometry analysis. In contrast, this project will use mass spectrometry and proteomic analysis to identify and quantify the proteins that are precipitated by the bait proteins. Tandem Mass Tag (TMT) labels – stable isotope labels that are covalently linked to peptides derived from bait protein immunoprecipitates and recognised by the mass spectrometer (Bantscheff *et al.*, 2007) – will be utilised to enable the quantification of differences between proteins precipitated with the bait proteins and a negative control.

In order to use the bait protein method to identify a Type 1 FCoV CER, a cell type bearing a Type 1 FCoV CER must be used for the proteomic screen. In the natural course of infection, FCoV spreads from the intestinal epithelium to monocytes and is conveyed systemically *via* a monocyte-associated viraemia (Kipar *et al.*, 1999a, Gunn-Moore *et al.*, 1998, Meli *et al.*, 2004, Desmarets *et al.*, 2016). Work by Van Hamme *et al.* (2007) demonstrated efficient binding and internalisation of a laboratory-adapted Type 1 FCoV strain by *ex vivo* feline monocytes, and a preliminary study suggested that viral pseudotypes expressing a Type 1 FCoV field strain S protein were able to infect *ex vivo* feline monocytes (Dye, 2006). These results indicate that feline monocytes are permissive to Type 1 FCoV entry and therefore likely carry a CER that the virus can utilise.

Intestinal epithelial cells, and particularly the mature cells situated on the villous tips, have been shown to contain FCoV antigen in cats naturally infected with the virus (Kipar *et al.*, 1998b), demonstrating that they are permissive to viral entry. This was confirmed when a feline intestinal epithelial cell line was developed and successfully infected with a field strain of Type 1 FCoV, albeit with low efficiency (Desmarets *et al.*, 2016). Unfortunately, this cell line is not available for this project, despite attempts to obtain it for research purposes.

Intestinal organoids are 3D cultures, derived from clusters of $Lgr5^+$ stem cells (also known as 'crypts') situated between intestinal villi, that self-organise into 'mini guts' containing all the cells present in the normal intestinal epithelium. Not only do these organoid cultures better recapitulate the organisation, functionality and heterogeneity of the intestine than cell lines, they also benefit from being self-propagating without the need for transformation (Ramani *et al.*, 2018). Intestinal organoids were first developed by Sato *et al.* (2009), who found that $Lgr5^+$ stem cells isolated from murine small intestine spontaneously developed into crypt-villus units. Intestinal organoids have since been established from a range of species including cats (Powell and Behnke, 2017), and have

been used to further our understanding of intestinal epithelial biology, as the basis of disease models and even to regenerate the intestinal epithelium (Date and Sato, 2015).

Just as researchers have struggled to propagate field strains of Type 1 FCoV *in vitro*, this was also the case for human noroviruses (HuNoVs); intestinal pathogens which for over 40 years could be not grown in cell culture until Ettayebi *et al.* (2016) demonstrated that human intestinal organoids support HuNoV replication. Similarly, human rotaviruses, which generally replicate poorly in transformed cell lines, grew successfully in human intestinal organoids and even induced physiological changes typical of rotavirus diarrhoea (Saxena *et al.*, 2016). Since Type 1 FCoV is also an intestinal pathogen, it is feasible that these breakthroughs could be repeated with feline intestinal organoids and Type 1 FCoV. If feline intestinal organoids could support *in vitro* propagation of field strains of Type 1 FCoV, this would circumvent the need to create a susceptible cell line.

4.1.1 Aims

The aims of this part of the project were to:

- Validate and optimise the bait protein method by using it to 'identify' the Type 2 FCoV CER, using cells known to express this receptor
- Identify a cell type bearing a Type 1 FCoV CER
- Use the bait protein method with this cell type to identify a Type 1 FCoV CER.

4.2 Results

4.2.1 Production and analysis of bait proteins

Plasmids encoding Type 1 (pT1) and Type 2 (pT2) FCoV bait proteins, produced previously by Quirke (2016), were transformed into bacteria and amplified in bacterial culture. Plasmid DNA was extracted using a midiprep procedure and analysed by restriction digestion (Figure 4.1). Agarose gel electrophoresis showed that the plasmids were of the expected size; pT1 and pT2 both showed bands of approximately 5,000 bp and 3,000 bp when digested, which corresponds with the expected sizes of the pCAGGS vector and inserts respectively. The plasmid DNA was purified and concentrated using ethanol precipitation, ready for transfection into mammalian cells for bait protein production.

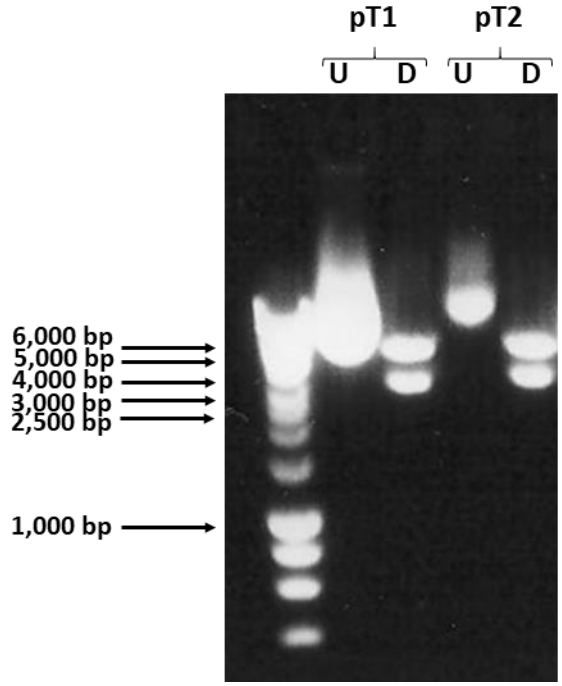

Figure 4.1. Analysis of bait protein-encoding plasmids by agarose gel electrophoresis. Plasmids encoding Type 1 (pT1) and Type 2 (pT2) FCoV bait proteins were amplified in bacterial culture, then the plasmid DNA was extracted and digested with appropriate restriction enzymes to liberate the inserts from their plasmids. The digestion products (D) were run alongside an aliquot of undigested plasmid (U) on a 1% agarose gel at 100 V for 40 minutes. The DNA bands were visualised under a UV transilluminator. The positions of relevant DNA mass markers are shown in bp.

HEK293T cells were transfected with pT1 and pT2, as well as pmaxGFP as a positive control and the empty parental pCAGGS vector as a negative control (obtained from Quirke (2016)). Using the transfection protocol developed by Quirke (2016), a transfection efficiency of approximately 50% (measured by the proportion of transfected cells expressing GFP) was observed. An experiment was set up to optimise the transfection procedure, whereby two variables – the source of HEK293T cells and age of Lipofectamine transfection reagent – were evaluated. The former variable was chosen because HEK293T cells ('HEK293T new') that had been clonally selected for optimal transfection efficiency in lentivirus packaging experiments were available in the laboratory, whereas the provenance of the cells used in the first transfection experiment in this project ('HEK293T old') was

unknown. The latter variable was chosen because the performance of Lipofectamine is only guaranteed for six months following purchase (Thermo-Fisher, 2015), and the batch used for the original transfection experiments ('Lipofectamine old') was over a year old. The new batch ('Lipofectamine new') was less than six months old. The two different cell lines were transfected with pmaxGFP, each using 'Lipofectamine old' and 'Lipofectamine new' and viewed using light and fluorescence microscopy (Figure 4.2). Transfection efficiency was similar (between 73% and 78%) under all conditions, except when 'Lipofectamine old' and 'HEK293T old' were used together and efficiency dropped to 52%. The efficiency of transfection was significantly associated with both HEK293T type and Lipofectamine age (p<0.001).

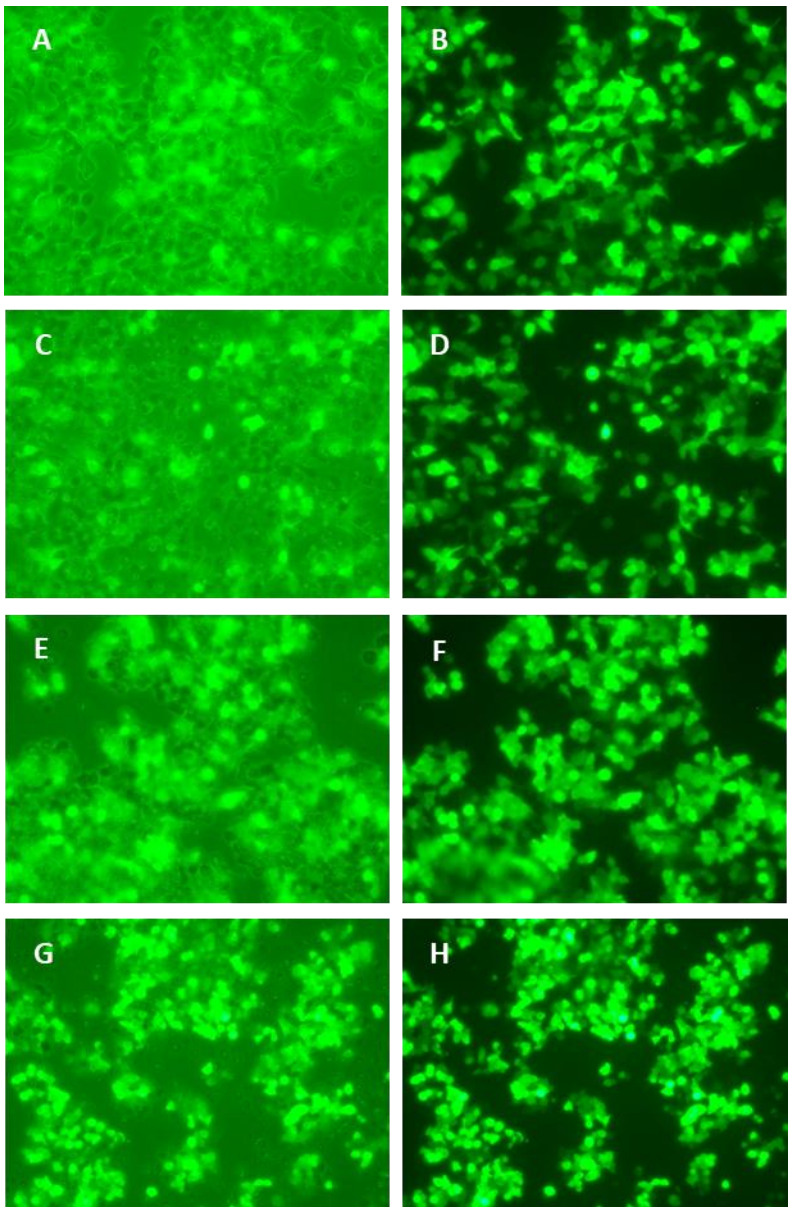

Figure 4.2. Examination of transfected cells to determine the impact on transfection efficiency of two variables. HEK293T cells, transfected with pmaxGFP, were viewed using light (left) and fluorescence (right) microscopy at 100x magnification, and the percentage of cells expressing GFP (green) was counted manually. The experiment was carried out in duplicate but only one set of results is shown.
A+B: 'HEK293T old', 'Lipofectamine new';
C+D: 'HEK293T old', 'Lipofectamine old';
E+F: 'HEK293T new', 'Lipofectamine new';
G+H: 'HEK293T new', 'Lipofectamine old'.

Following transfection with pT1, pT2 and empty pCAGGS using the optimised conditions, transfected cells were cultured for 72 hours (identified as the optimal culture length by Quirke (2016)) and the cell culture supernatants were harvested. Some of the culture supernatant containing each bait protein was purified by affinity chromatography. The raw and purified supernatants were examined by SDS-PAGE and Coomassie Brilliant Blue staining or western blot (Figure 4.3), which showed that bait proteins of the correct sizes (140 kDa for Type 1 and 150 kDa for Type 2) were produced and secreted into the culture supernatant, and were still present in the purified eluate.

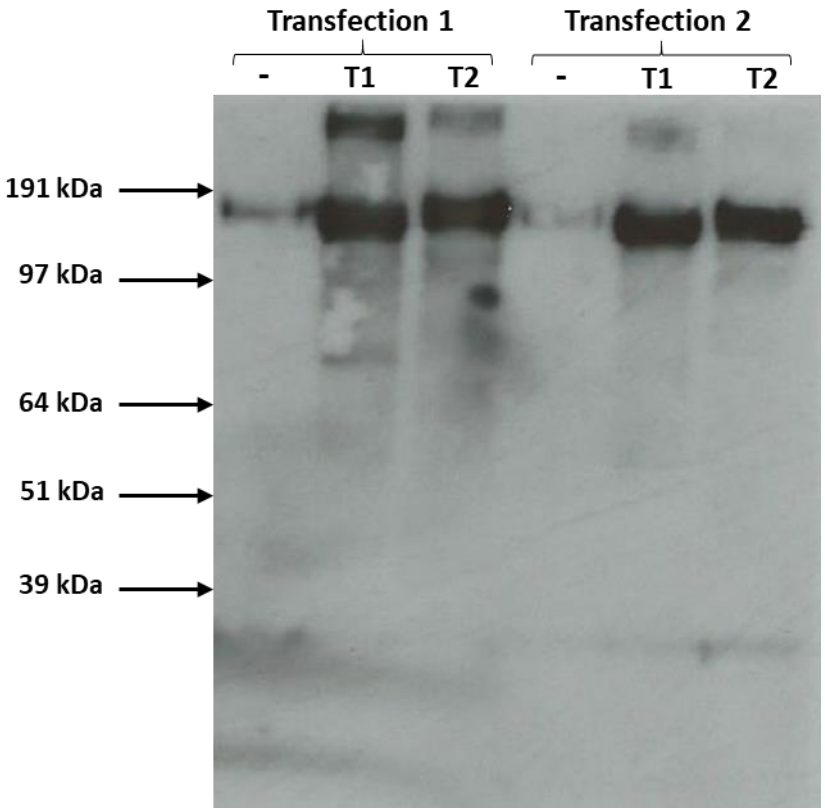

Figure 4.3. Demonstration of the presence of bait protein in transfected cell culture supernatant. Proteins in equal volumes of raw cell culture supernatant, produced from approximately equal numbers of cells transfected with pT1 (T1), pT2 (T2) or empty pCAGGS vector (-), were denatured and run on a 4-12% SDS-PAGE Bis-Tris mini gel. Proteins in the gel were transferred onto a membrane and this was probed with anti-human IgG Fc antibody. The membrane was then incubated with a horseradish peroxidase-labelled secondary antibody, prior to application of a chemiluminescent substrate and exposure to x-ray film. Bait proteins produced in two separate transfections (Transfection 1 and Transfection 2) are shown. The positions of relevant molecular mass markers are shown in kDa.

4.2.2 Validation of the bait protein method

In order to validate and optimise the bait protein method as a way of identifying a Type 1 FCoV CER, the method was first used to 'identify' the Type 2 FCoV CER on CrFK cells, which are known to carry the Type 2 FCoV CER, fAPN.

An immunofluorescence assay was carried out to screen CrFK cells with the bait proteins (Figure 4.4). This was done to check whether Type 2 bait protein was capable of recognising fAPN on the surface of CrFK cells, and to compare the performance of neat and purified and concentrated Type 2 bait protein. Where the bait protein was used as neat culture supernatant, it was diluted 1 in 2 in blocking buffer before use in the assay. Although this dilution was simply chosen as a starting point, the assays worked well, so no further optimisation of the bait protein dilution was carried out. By visualising different dilutions of neat and purified bait protein alongside each other with SDS-PAGE and Coomassie Blue staining (data not shown), neat bait protein at a dilution of 1 in 2 was calculated to be equivalent to purified bait protein at a dilution of 1 in 14.

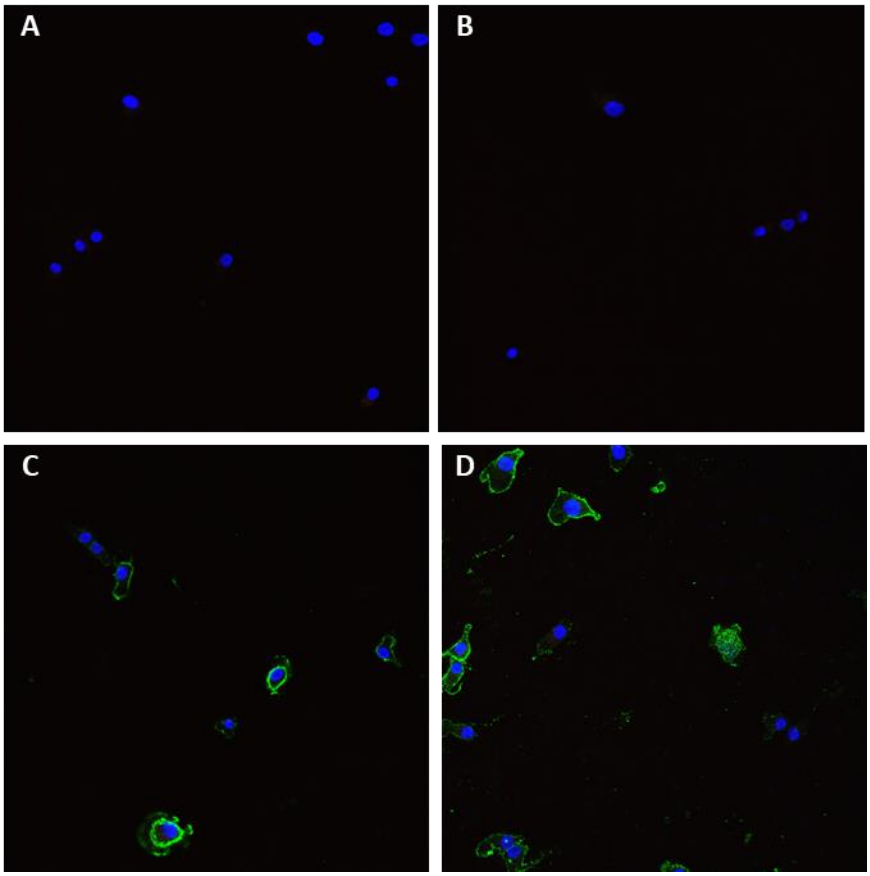

Figure 4.4. Examination of CrFK cells by immunofluorescence assay for bait protein recognition. CrFK cells were grown on glass coverslips then fixed without permeabilisation. The cells were incubated with neat Type 1 (1 in 2; B), neat Type 2 (1 in 2; C), purified Type 2 (1 in 14; D) or no (A) bait protein, followed by a DyLight 488-conjugated secondary antibody (green). Nuclei were stained with DAPI (blue). Cells were then viewed using a confocal laser scanning microscope with a 40x oil immersion objective.

The immunofluorescence assay demonstrated staining with the Type 2 but not Type 1 bait protein, indicating a specific interaction between the Type 2 bait protein and CrFK cells. This staining

occurred with both neat and purified Type 2 bait protein, with little visible difference between the two. From this point of the project onwards, only neat bait protein was used.

The specific interaction between CrFK cells and Type 2 bait protein was confirmed using flow cytometry (Figure 4.5), which demonstrated a clear divide between cells incubated with Type 2 and no bait protein.

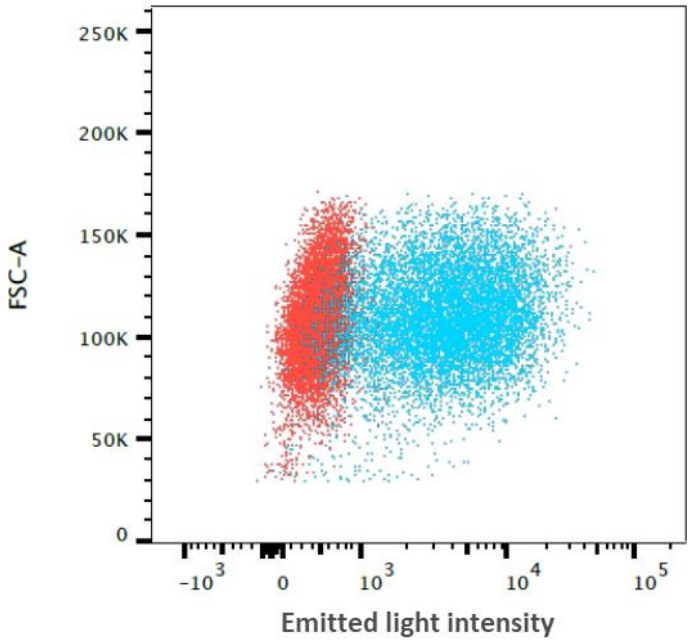

Figure 4.5. Examination of CrFK cells by flow cytometry for bait protein recognition. A CrFK cell suspension was made using ACCUTASE™ cell detachment solution. Cells were incubated with Type 2 or no bait protein, followed by a CF 633-conjugated secondary antibody. Cells were analysed on an LSR II flow cytometer, using a 633 nm laser for excitation and a 660/20 band pass filter for detection. The plot displays intensity of light emitted from CrFK cells incubated with Type 2 (blue dots) and no (red dots) bait protein and is gated to include only live, single cells. FSC-A: forward scatter (a measure of cell size).

The next step was to determine by immunoprecipitation, mass spectrometry and proteomics what the Type 2 bait protein recognised on CrFK cells. Following the protocol that Raj *et al.* (2013) used to identify the CER for MERS-CoV, lysates were prepared from CrFK cells using a detergent (n-Dodecyl ß-D-maltopyranoside; DDM) designed to isolate cell surface proteins. The lysates were incubated with the Type 1 and 2 bait proteins, plus, as a negative control, culture supernatant not containing bait protein. This in turn was incubated with Protein A resin, which binds the IgG Fc portion of the bait proteins, coprecipitating anything which the bait proteins have bound. A proportion of the proteins eluted from the Protein A resin were analysed by SDS-PAGE and Coomassie Blue staining (Figure 4.6), which showed bands for the Type 1 and 2 bait proteins but no other visible difference between the three conditions. A 109 kDa band representing fAPN was not detectable in the Type 2 bait protein immunoprecipitate, but this may have simply meant that the amount of fAPN was below the level of detection for Coomassie blue staining. Very strong bands were present at approximately 160 kDa under all three conditions, suggesting a strong, non-specific interaction between Protein A and a component of the eluates. Using the same protocol this protein band was also recognised by

Quirke (2016), who analysed the protein by mass spectrometry and identified it as feline myosin-9. A step to pre-clear myosin-9 from the CrFK cell lysate and culture supernatants was considered, but Quirke (2016) found this to have little impact on the amount of myosin-9 in the final eluates. This step was therefore omitted.

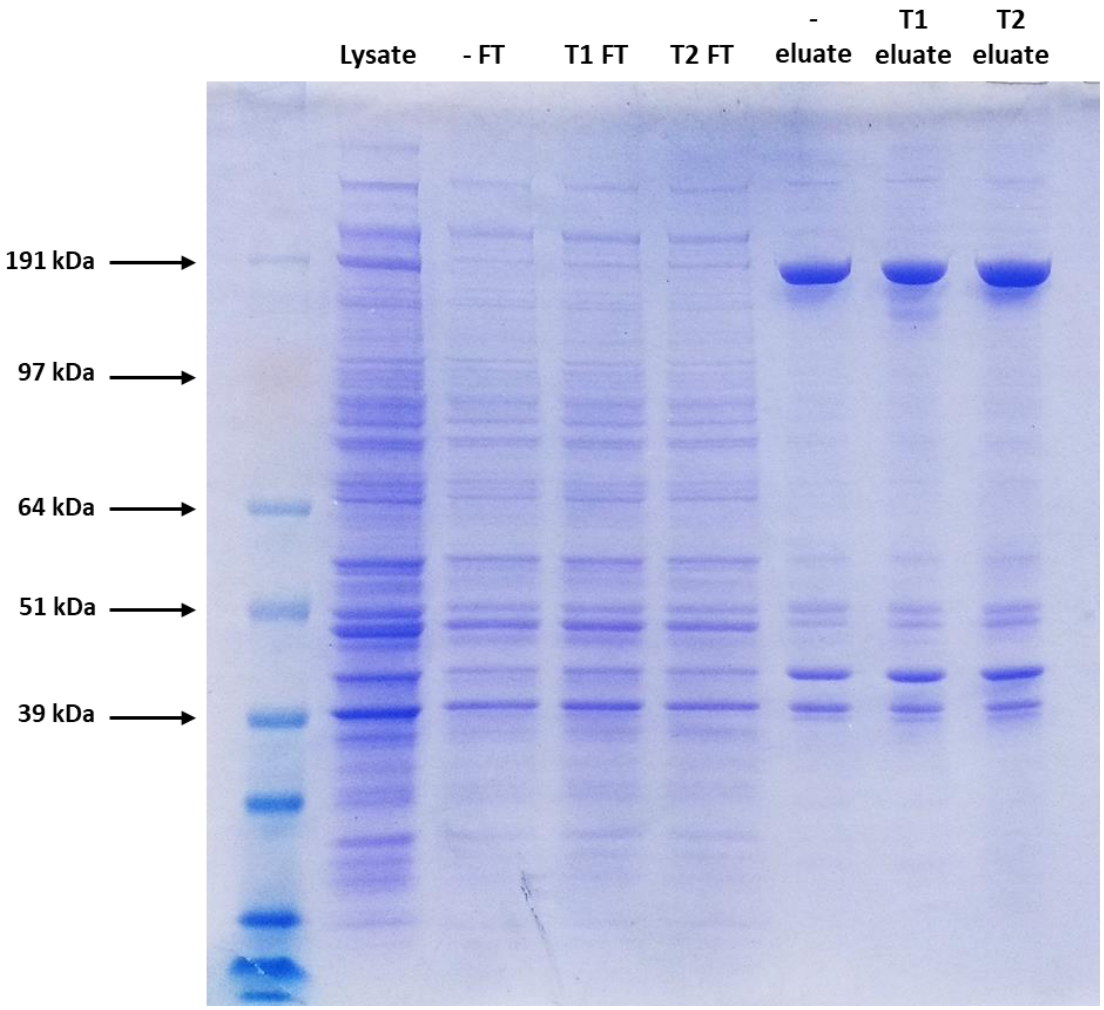

Figure 4.6. Analysis of the proteins present at different stages of an immunoprecipitation experiment with CrFK cells and the bait proteins. CrFK cells were lysed and membrane proteins solubilised using a detergent designed specifically to isolate cell surface proteins (Lysate). The lysate was incubated with an equal volume of Type 1 (T1), Type 2 (T2) or no (-) bait protein, and these in turn were incubated with Protein A resin, which binds the IgG part of the bait proteins. Lysate, proteins that did not bind to Protein A resin (flow through; FT) and proteins that bound to and were eluted from Protein A resin (eluate) were denatured and run on a 4-12% SDS-PAGE Bis-Tris mini gel, followed by Coomassie Brilliant Blue staining. The positions of relevant molecular mass markers are shown in kDa.

Despite the issue with myosin-9, it was decided to continue with the procedure because this project used quantitative mass spectrometry, so any protein enriched under all conditions could be excluded from the analysis. Immunoprecipitated proteins were digested in solution, and the resulting peptides TMT-labelled and identified through mass spectrometry. Only 18 feline proteins, none of which were fAPN, were identified, despite the detection of more than 18 proteins by SDS-PAGE and Coomassie Blue staining.

The experiment was repeated as described above and similarly resulted in an extremely limited protein list. At this point it was unknown whether the problem lay with the immunoprecipitation

protocol, peptide preparation (i.e. in-solution digestion or TMT-labelling) or the feline protein database. To troubleshoot, a CrFK whole cell lysate was analysed by mass spectrometry, without TMT-labelling. Additionally, in-gel (as opposed to in-solution) digestion was carried out to remove any residual DDM from the samples; since it is not a commonly used detergent in immunoprecipitation experiments, DDM may have been causing an issue. Over 2,000 feline proteins were identified (Appendix D), indicating that the problem was with immunoprecipitation or peptide preparation as opposed to the feline database. Next, this in-gel digestion and non-TMT approach was carried out on proteins immunoprecipitated from a CrFK cell lysate by the bait proteins. Over 1,000 proteins were identified (Appendix D), indicating that immunoprecipitation was not the issue and rather the problem lay with peptide preparation. Though not quantitative, this experiment identified fAPN as one of the 20 most abundant proteins, compared to the negative control in which fAPN was not identified at all.

Finally, proteins obtained through immunoprecipitation by the bait proteins underwent in-gel digestion and TMT-labelling. Over 700 proteins were identified through mass spectrometry (Appendix D), and fAPN was identified as the second hit for the immunoprecipitation with the Type 2 bait protein. In contrast, the top five proteins for the immunoprecipitation with the Type 1 bait protein did not include fAPN, and instead comprised proteins that are not known to localise at the cell surface or function as receptors (Tables 4.1 and 4.2). The top hit for the immunoprecipitation with the Type 2 bait protein, putative small glutamine-rich tetratricopeptide repeat-containing protein alpha (SGTA), was the fourth most enriched by Type 1 bait protein, suggesting a common interaction with both serotypes.

Having optimised the immunoprecipitation, mass spectrometry and proteomics procedure and ‘identified’ fAPN on CrFK cells using the Type 2 bait protein, the procedure was now ready to be used for the identification of the Type 1 FCoV CER.

Table 4.1. The top five proteins immunoprecipitated from CrFK cells by the Type 1 bait protein. Proteins were identified following in-gel digestion, TMT-labelling and mass spectrometry of the immunoprecipitates, then sorted based on the log2 fold change in their abundance compared with the control immunoprecipitation.

Uniprot accession number	Protein description	Log2 fold change over control
M3WVF5	Putative cytoplasmic dynein 1 light intermediate chain 2	6.50
M3W0U2	Putative serine/threonine-protein kinase Nek3	4.85
P07405	Haemoglobin subunit alpha	4.13
M3WN94	Putative small glutamine-rich tetratricopeptide repeat-containing protein alpha	4.07
M3W6G9	Putative nuclear prelamin A recognition factor	4.00

Table 4.2. The top five proteins immunoprecipitated from CrFK cells by the Type 2 bait protein. Proteins were identified following in-gel digestion, TMT-labelling and mass spectrometry of the immunoprecipitates, then sorted based on the log2 fold change in their abundance compared with the control immunoprecipitation.

Uniprot accession number	Protein description	Log2 fold change over control
M3WN94	Putative small glutamine-rich tetratricopeptide repeat-containing protein alpha	4.08
A0A337SS27	Aminopeptidase N	3.62
A0A2I2V231	TOG domain-containing protein	2.59
M3XFZ3	Annexin	2.41
M3X0N0	Isocitrate dehydrogenase	2.38

4.2.3 Screening cell lines for a Type 1 FCoV CER

On validating and optimising the bait protein method, work began to identify a cell type expressing a Type 1 FCoV CER. Two cell lines that had not been previously tested for their permissibility to Type 1 FCoV infection, DH82 (a canine macrophage-like line) and FE-A (a feline embryonic line), were analysed. Though a canine cell line, DH82 was chosen because FCoV has a tropism for macrophages and, owing to the evolutionary relatedness of cats and dogs and the origin of Type 2 FCoV, interspecific circulation of FCoV in dogs is plausible (Le Poder, 2011). As DH82 is a macrophage-like cell line and macrophages are known to express Fc receptors, human IgG Fc was used to control for Fc receptor binding. To determine a concentration of human IgG Fc equivalent to the bait proteins in culture supernatant, three dilutions of IgG Fc were analysed alongside the bait proteins by SDS-PAGE and western blotting (Figure 4.7). Bands of the expected size were detected for human IgG Fc (32 kDa), Type 1 bait protein (140 kDa) and Type 2 bait protein (150 kDa). It is unknown what the additional non-specific bands represented. Neat bait protein was estimated to contain the same amount of protein as IgG Fc diluted 1:500, which was equivalent to 4.6 µg/ml.

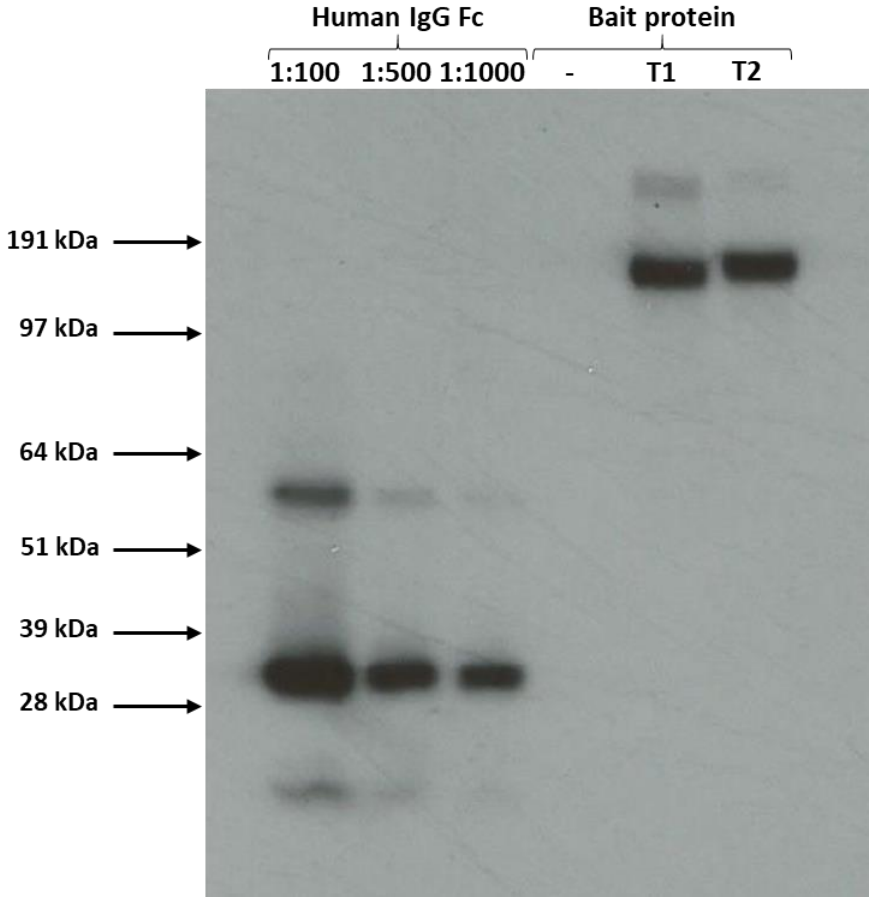

Figure 4.7. Analysis by western blot of the concentration of protein in human IgG Fc compared with culture supernatant containing the Type 1 and 2 bait proteins. Human IgG Fc at various dilutions, neat Type 1 bait protein-containing culture supernatant (T1), neat Type 2 bait protein-containing culture supernatant (T2) or neat culture supernatant containing no bait protein (-) was denatured and loaded onto a 4-12% SDS-PAGE Bis-Tris mini gel. Proteins in the gel were transferred onto a membrane and this was probed with anti-human IgG Fc antibody. The membrane was then incubated with a horseradish peroxidase-labelled secondary antibody, prior to application of a chemiluminescent substrate and exposure to x-ray film. The positions of relevant molecular mass markers are shown in kDa.

On determining the correct dilution of IgG Fc to use, an immunofluorescence experiment was carried out using the bait proteins and IgG Fc as a negative control with DH82 and FE-A cells (Figure 4.8). The results demonstrated specific staining of DH82 cells with the Type 2 but not the Type 1 bait protein, and no staining of FE-A cells with either bait protein. This suggested that something on the surface of DH82 cells was recognised by the Type 2 bait protein, but that neither DH82 nor FE-A cells carry a Type 1 FCoV CER.

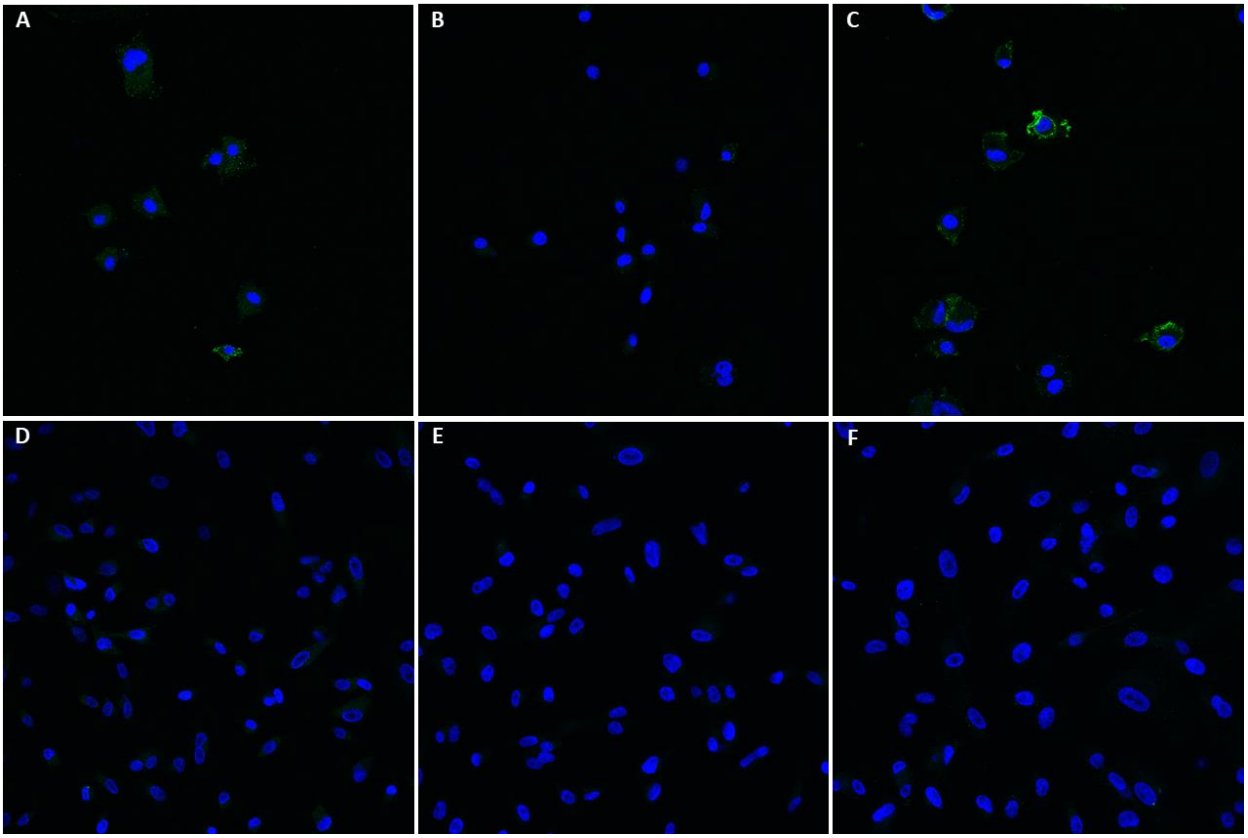

Figure 4.8. Examination of DH82 and FE-A cells by immunofluorescence assay for bait protein recognition. Cells were grown on glass coverslips then fixed without permeabilisation. The cells were incubated with either the Type 1 bait protein, Type 2 bait protein or human IgG Fc – negative control, followed by a DyLight 488-conjugated secondary antibody (green). Nuclei were stained with DAPI (blue). Cells were then viewed using a confocal laser scanning microscope with a 40x oil immersion objective. These results are typical of two replicates.
A: DH82, IgG Fc – negative control; B: DH82, Type 1 bait protein; C: DH82, Type 2 bait protein;
D: FE-A, IgG Fc – negative control; E: FE-A, Type 1 bait protein; F: FE-A, Type 2 bait protein.

4.2.4 Screening feline PBMC for a Type 1 FCoV CER

Feline PBMC were chosen next to be screened for a Type 1 FCoV CER because the virus infects monocytes – a component of PBMC – in the natural course of a Type 1 infection (Kipar and Meli, 2014). The PBMC were prepared from individual whole blood samples in EDTA using a density gradient medium method. All but one of the blood samples were obtained from the collection established for the ELISpot analysis (see Chapter 5), where a free FCoV antibody titre test was offered in exchange for a feline blood sample. Samples from 16 cats were not used for ELISpot analysis due to insufficient time, so the PBMC were frozen down and used in this part of the project instead. These 16 cats were all from the UK and comprised four Ragdolls, four Maine Coons, four British Shorthair, one Sphinx, one Exotic Shorthair, one Burmese and one Persian. Their ages ranged from five months to nine years, with over half of cats three years or less. Cats of all sexes were represented, with over half unneutered. The remaining blood sample was obtained *post mortem* from an adult, female neutered Domestic Shorthair cat, who was euthanased for reasons unrelated to this study.

PBMC from 10 cats were screened with the Type 1 and 2 bait proteins using an immunofluorescence assay, with human IgG Fc used again as a control for Fc receptor binding (a typical example is shown in Figure 4.9).

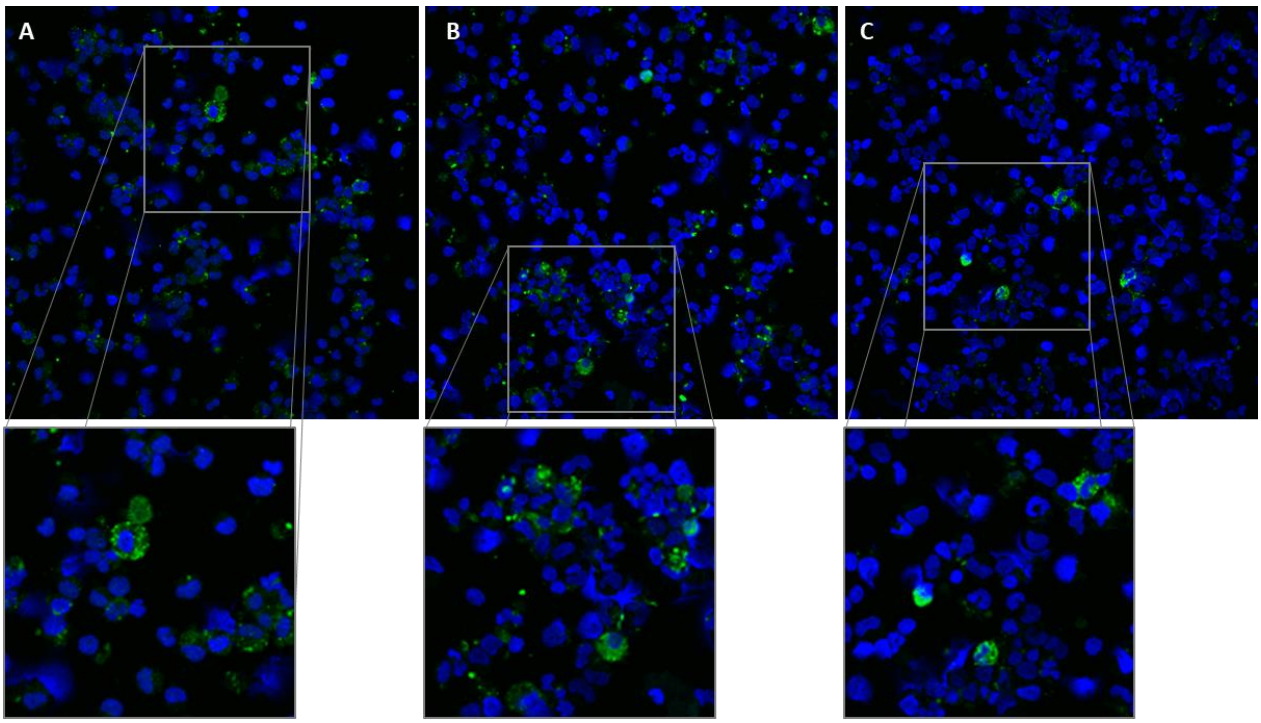

Figure 4.9. Examination of feline PBMC by immunofluorescence assay for bait protein recognition. PBMC were applied to glass coverslips by cytocentrifugation, then fixed without permeabilisation. The cells were incubated with the the Type 1 bait protein, Type 2 bait protein or a human IgG Fc – negative control. The cells were then incubated with a DyLight 488-conjugated secondary antibody (green). Nuclei were stained with DAPI (blue). Cells were then viewed using a confocal laser scanning microscope with a 40x oil immersion objective. These results are typical of five replicates. Close up views of an area of each image, indicated by the grey box, are displayed beneath their corresponding images.
A: IgG Fc – negative control; B: Type 1 bait protein; C: Type 2 bait protein.

The assay showed staining with the Type 1 and 2 bait proteins, but the pattern was irregular and the intensity variable. There was also some staining with the IgG Fc fragment, indicating that the bait protein interaction with PBMC did not represent specific recognition of a CER. Flow cytometry was carried out on PBMC from two cats to see if a specific interaction between a population of the cells and the Type 1 bait protein could be identified using a different method, but on both occasions no difference was observed between feline PBMC incubated with the negative control (IgG Fc) and Type 1 bait protein (Figure 4.10).

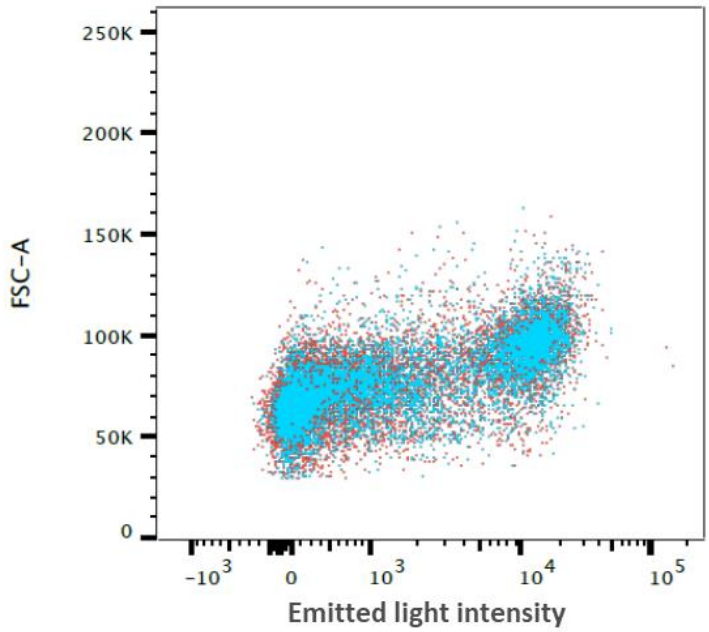

Figure 4.10. Examination of feline PBMC by flow cytometry for bait protein recognition. A suspension of ≥200,000 live PBMC was incubated with 100 µl neat Type 1 bait protein or human IgG Fc – negative control, followed by a CF 633-conjugated secondary antibody. Cells were analysed on an LSR II flow cytometer, using a 633 nm laser for excitation and a 660/20 band pass filter for detection. The plot displays intensity of light emitted from cells incubated with the Type 1 (blue dots) and no (red dots) bait protein and is gated to include only live, single cells. FSC-A: forward scatter (a measure of cell size). These results are typical of two replicates.

Other studies demonstrating monocyte entry and infection with FCoV did so after the cells adhered to a tissue culture surface and underwent a period of differentiation (Van Hamme *et al.*, 2007, Dewerchin *et al.*, 2005), so this approach was taken next. Monocytes from the remaining five PBMC samples were allowed to adhere to glass coverslips for 24 hours prior to washing away unadhered cells, which is recognised as an effective way of obtaining a cell population mostly consisting of monocytes (Dewerchin *et al.*, 2005). The monocytes were cultured for a further 4 days or until they began to take on a macrophage-like appearance: becoming larger, in some cases becoming elongated and often developing processes (data not shown). The resulting monocyte-derived cells were screened with the Type 1 and 2 bait proteins, and human IgG Fc as a negative control, in an immunofluorescence assay (Figure 4.11).

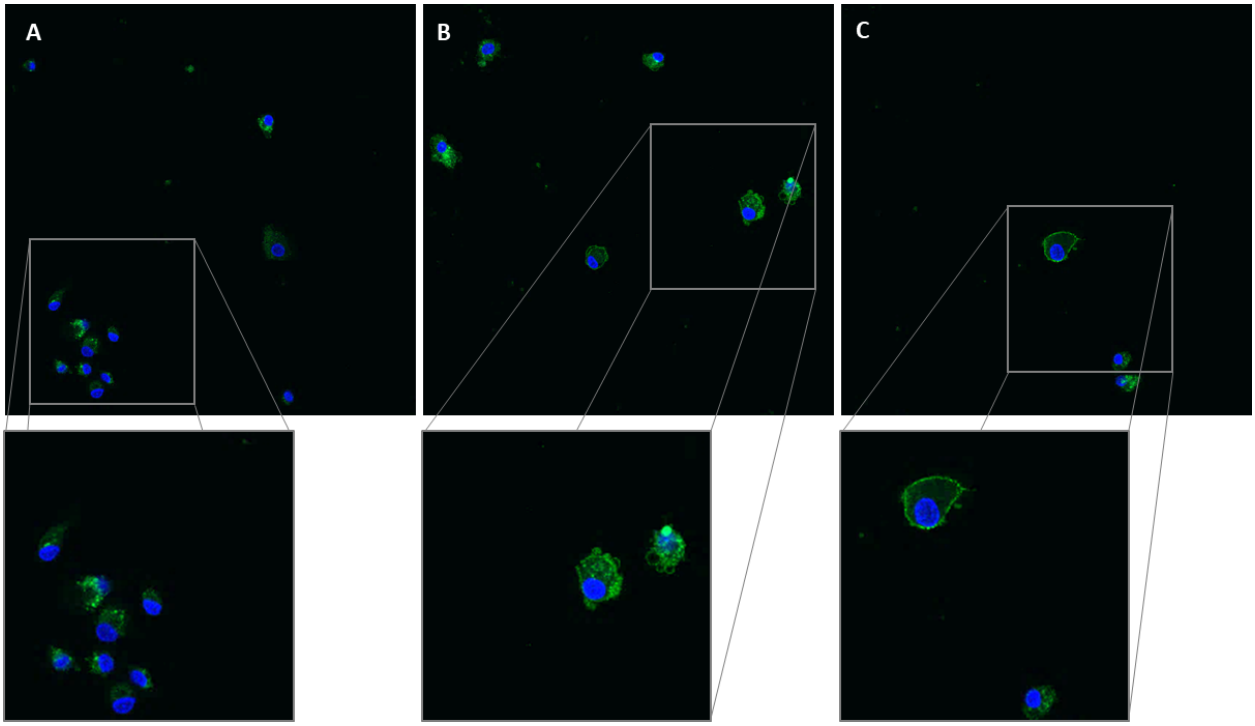

Figure 4.11. Examination of feline monocyte-derived cells by immunofluorescence assay for bait protein recognition. Monocytes were selected from PBMC by adherence to glass coverslips following culture for 24 hours. These monocytes were cultured for a further four days, then fixed without permeabilisation. The cells were incubated with the Type 1 bait protein, Type 2 bait protein or human IgG Fc – negative control, followed by a DyLight 488-conjugated secondary antibody (green). Nuclei were stained with DAPI (blue). Cells were then viewed using a confocal laser scanning microscope with a 40x oil immersion objective. These results are typical of five replicates. Close up views of an area of each image, indicated by the grey box, are displayed beneath their corresponding images.
A: IgG Fc – negative control; B: Type 1 bait protein; C: Type 2 bait protein.

The results appear to demonstrate cell surface staining of a proportion of monocyte-derived cells with both the Type 2 and Type 1 bait proteins, but this was difficult to differentiate from the green fluorescence detected in the negative control condition and was not reliably repeatable.

At this point no more PBMC were available for this study, and it was decided not to collect more because the flow cytometry experiments showed no specific staining of the PBMC with Type 1 bait protein, and the results of the immunofluorescence assays with differentiated monocytes were inconclusive.

4.2.5 Establishing feline intestinal organoid cultures

As the experiments investigating the presence of the Type I FCoV CER on PBMC were inconclusive, it was decided to focus next on feline intestinal epithelial cells. This cell type was chosen since it is a natural target for infection with Type 1 FCoV (Kipar and Meli, 2014). Initial attempts to isolate feline intestinal epithelial cells were made following a slightly modified protocol of Desmarets *et al.* (2013), where intestinal tissue was washed thoroughly of mucus and debris then enzymes were used to dissociate the epithelium from underlying structures. The epithelium was mechanically removed, digested further with enzymes to produce a cell suspension and pelleted. This procedure yielded a loose and mucinous cell pellet which, on cytological analysis, was found to contain low cell numbers, mostly poor cell preservation and moderate to high amounts of mucus and microbes (data not shown). The procedure was repeated, this time using a cell strainer to remove any undigested pieces of tissue and mucus from the cell suspension. The resulting suspension was seeded into a cell culture flask, but no cells were recovered. This was carried out once more with the same result. Due to the difficulties experienced with this method and the relative advantages of organoids compared to immortalised cell lines (described in section 4.1), the focus of the investigation changed to the establishment of feline intestinal organoid cultures.

Initially murine intestinal organoids, obtained as *in vitro* cultures from Dr R. Jenkinson, University of Bristol, were used to practice organoid subculture technique (Figure 4.12). L-WRN conditioned medium, made in-house following the protocol of Miyoshi and Stappenbeck (2013), was also tested at this point with a view to using it for feline intestinal organoid cultivation as reported by Powell and Behnke (2017). The murine organoids were grown in two media in parallel: conditioned medium and the commercially available IntestiCult organoid growth medium mouse (IC mouse). Initially, the organoids in IC mouse grew much more robustly than those in conditioned medium, forming large, budding structures within a few days of passage (Figure 4.13). The organoids in conditioned medium remained small and did not bud. However, after two passages the growth of organoids in both media was comparable, suggesting that the organoids had adapted to the conditioned medium.

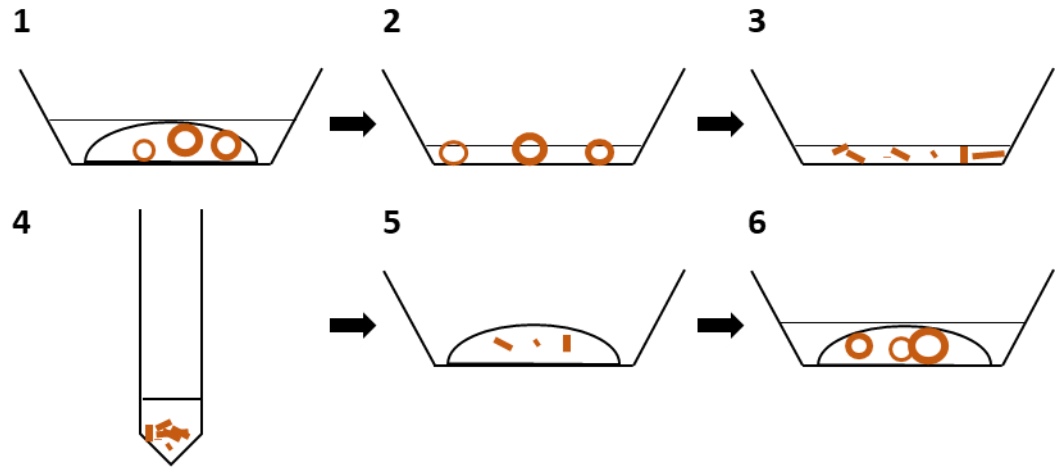

Figure 4.12. Schematic to show the process for subculturing organoid cultures. Each step shows a side view of one well of a cell culture plate, except for step 4 which shows a side view of a conical centrifuge tube.
1: Organoids, suspended in Matrigel matrix and submerged in medium, are ready to be subcultured.
2: The Matrigel matrix is dissociated by placing the plate on ice or mechanically disrupting the matrix by pipetting.
3: Once the Matrigel matrix has turned to liquid, the organoids can be fragmented by enzymatic digestion and/or mechanical disruption by pipetting.
4: The organoid fragments are collected into a centrifuge tube and pelleted by centrifugation.
5: The organoid fragments are resuspended in Matrigel matrix and an aliquot is seeded into a new well.
6: The organoid fragments, suspended in Matrigel matrix and submerged in medium, reform into spheroids and grow larger.

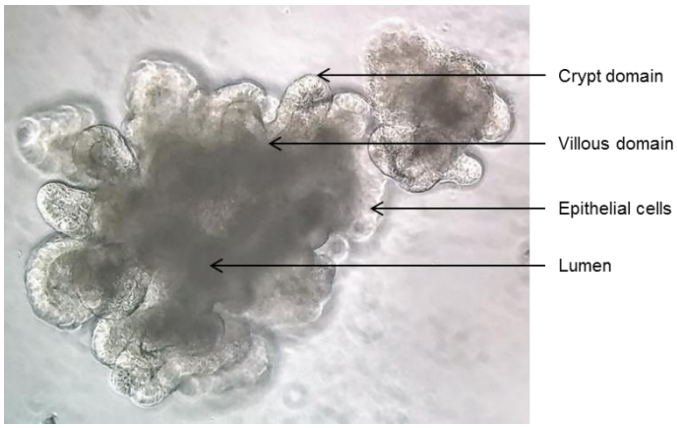

Figure 4.13. Examination of a large, budding murine intestinal organoid. This organoid was cultured in Matrigel matrix, submerged in IC mouse, for around 5 days following passage, then imaged under a light microscope at 200x magnification.

Having had some practice handling the murine intestinal organoids and ascertained that the conditioned medium was fit for purpose, the next step was to establish feline intestinal organoid cultures. Feline organoids were first obtained from Dr M. Behnke, Louisiana State University, as frozen cultures (Table 4.3). The first feline organoids that were thawed and cultured were from Cat 11, Passage 10 (C11P10), and these were grown in conditioned medium containing nicotinamide, Y-27632 and SB-43154 (as instructed by Dr M. Behnke). The published report describing feline organoid culture states that cat organoids cease to expand at around passage 10 and undergo growth arrest at passage 13 to 18 (Powell and Behnke, 2017), but in personal correspondence Dr M. Behnke explained that this problem has since been overcome by adding a high concentration of nicotinamide to the conditioned medium. This high concentration of nicotinamide was used for feline intestinal organoid culture throughout this project.

Table 4.3. Summary of the feline organoids available as frozen cultures from Dr M. Behnke. Ten frozen vials were received in total, containing organoids from three cats at different passage numbers.

Cat ID	Passage numbers available	Cat age
11	6, 10, 65, 89	Adult
14	4, 7	Fetus
19	4, 16, 23, 28	Adult

Though lots of cellular material was present when C11P10 was plated, only a single organoid survived, and this died after the first passage.

The next organoids that were thawed and cultured were from Cat 19, Passage 16 (C19P16), and these were grown in conditioned medium, IC mouse and IntestiCult organoid growth medium human (IC Human). Those in conditioned medium were initially viable but never expanded and died after the first passage. Those in IC mouse and IC human survived and grew very well, with those in IC human becoming particularly large and cystic after the second passage. However, following the third passage the organoids in both IC mouse and IC human remained very small and died after around 7 days.

Organoids from Cat 14, Passage 7 (C14P7) were thawed and cultured, and these were initially split into conditioned medium and IC human. In parallel, murine organoids were thawed and cultured in conditioned medium to re-test its ability to support organoid growth. The murine organoids grew very robustly in conditioned medium, showing again that the medium could support organoid growth. This time the feline organoids grew better in conditioned medium than in IC human, with more viable organoids reaching larger sizes. Following the first passage, all the feline organoids were grown in conditioned medium. C14P7 survived for seven passages, with growth slowly declining from around the third passage. The organoids that grew were large and cystic with no budding, as is typical for feline intestinal organoids (Powell and Behnke, 2017).

Though C14P7 survived for longer than any of the previous attempts, they still ceased to expand far sooner than reported by Powell and Behnke (2017). To eliminate the possibility that the reagents could be causing a problem, Dr M. Behnke kindly donated aliquots of all the reagents used for organoid culture in his laboratory. Organoids from Cat 14 Passage 4 were split into two, with half cultured using in-house reagents and the other half cultured using reagents from Dr M. Behnke. In both cases, growth was similarly sluggish and only a few passages were achieved before the cultures began to die off.

Considering the problems experienced recovering frozen organoid cultures and propagating them beyond a few passages, work began to isolate crypts from feline intestinal tissue and cultivate organoids from them. Intestinal tissue was obtained from three adult Domestic Shorthair cats who

were euthanased at a local animal shelter for reasons unrelated to this project. None of the cats were known to have intestinal disease, and their FCoV status was undefined. The tissue was collected no longer than 30 minutes after death and processed within an hour following the protocol of Powell and Behnke (2017), which involved cleaning, mincing and enzymatically digesting the tissue, then straining and centrifuging the resulting cell suspension to enrich for crypts. The remainder of the protocol is illustrated in Figure 4.12, steps 5 and 6. Organoids were grown in conditioned medium supplemented with nicotinamide, Y-27632 and SB-43154. Figure 4.14 depicts the cultivation of feline intestinal organoids from crypts.

Organoids from all three cats grew readily from the isolated crypts. Vigorous cultures were established by day 7 in all cases, organoids were passaged every 3-7 days and 7-9 passages were achieved before organoids began growth arrest. Most organoids displayed a cystic phenotype, but in some cases there was a mixed phenotype within the same culture of cystic and budding (Figure 4.15). Cells with a mesenchymal morphology were observed to co-culture with the organoids (indicated by a white arrow; Figure 4.16) until around passage 7, at which point they died off.

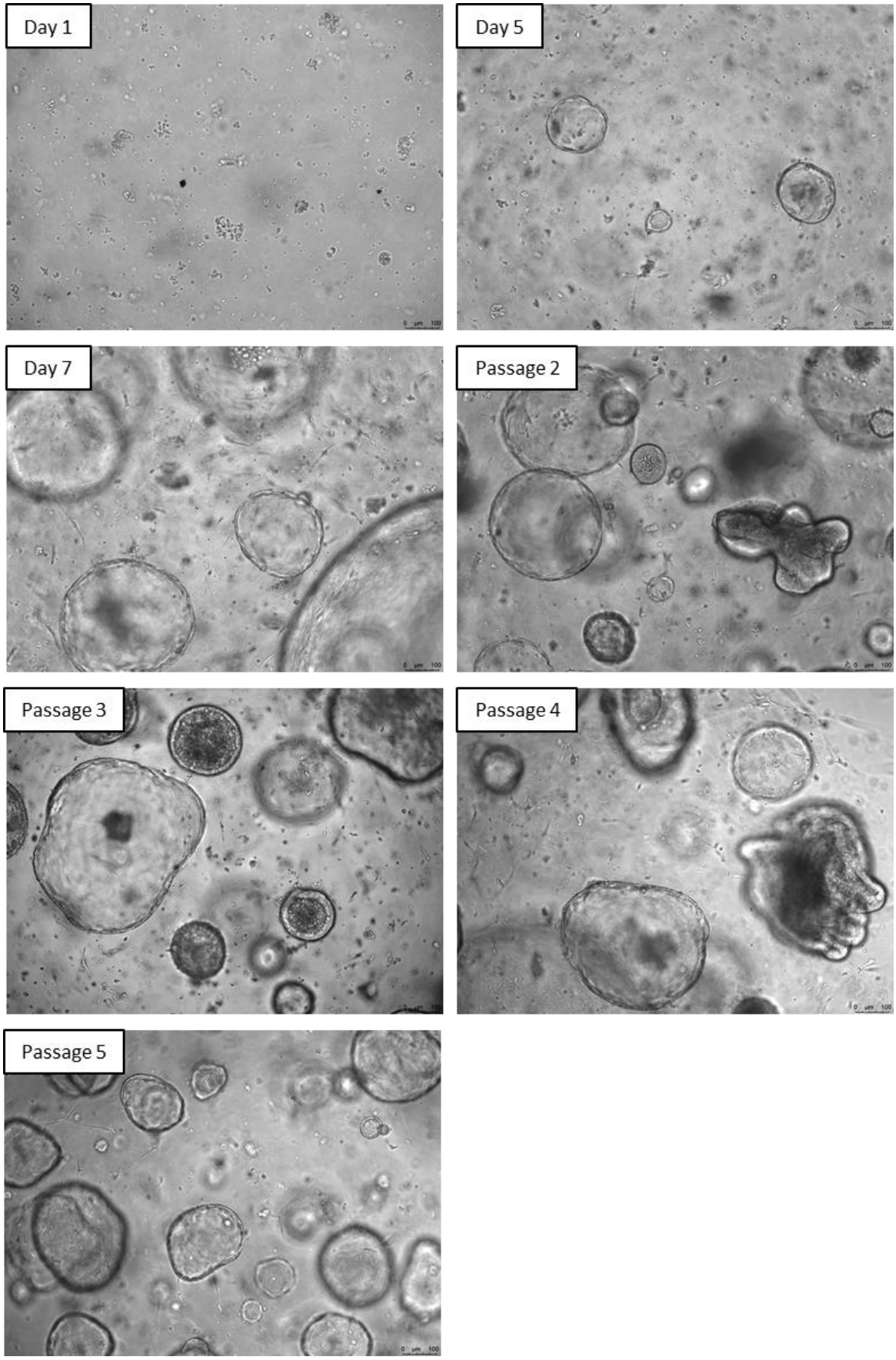

Figure 4.14. The cultivation of feline intestinal organoids from intestinal crypts. Organoids, derived in-house from feline intestinal tissue, were cultured in Matrigel matrix, submerged in conditioned medium, and imaged under a light microscope at 100x magnification at various time points, as indicated. Organoids were passaged every 3-7 days.

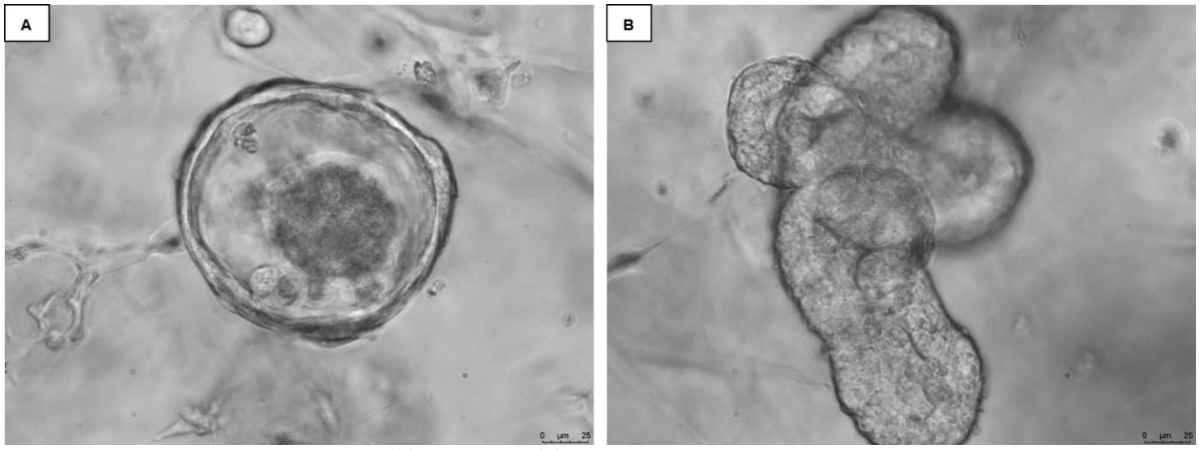

Figure 4.15. Examination of a typical cystic (A) and budding (B) feline intestinal organoid. Organoids, derived in-house from feline intestinal tissue, were cultured in Matrigel matrix, submerged in conditioned medium, for around 5 days following passage, then imaged under a light microscope at 400x magnification.

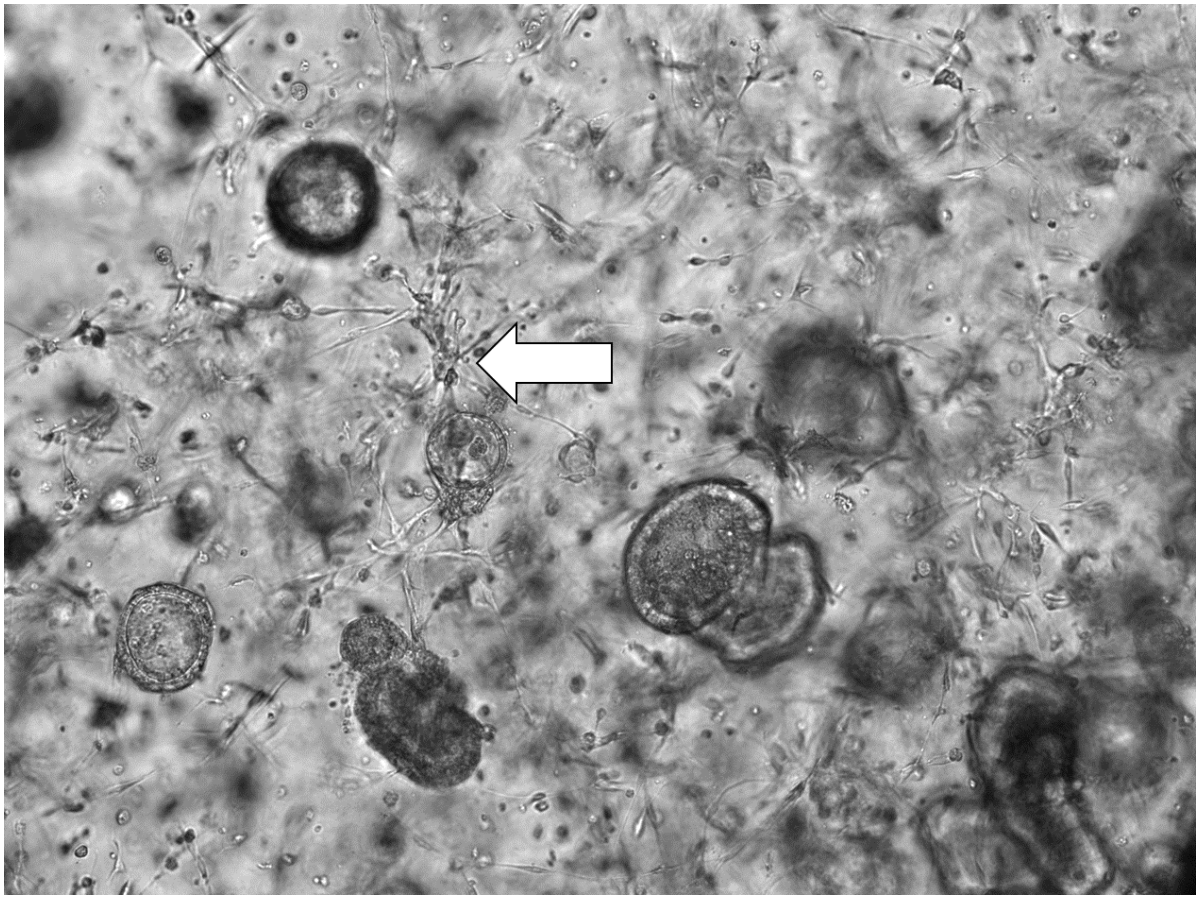

Figure 4.16. Examination of a mixed culture containing feline intestinal organoids and spindle cells. Organoids, derived in-house from feline intestinal tissue, were cultured in Matrigel matrix, submerged in conditioned medium, for around 5 days following passage 4, then imaged under a light microscope at 100x magnification. Many mesenchymal cells are present in this image, but a typical example of a group of such cells is indicated with a white arrow.

4.2.6 Screening feline intestinal organoids for a Type 1 FCoV CER

Though feline intestinal organoid cultures could not be propagated beyond nine passages, it was decided to go ahead and screen them for a Type 1 FCoV CER, rather than spend more time optimising their cultivation. Initially, the organoids were screened for bait protein recognition in an immunofluorescence assay that involved fixing and staining the organoids *in situ*, embedded in Matrigel matrix. The results of this assay were extremely difficult to interpret with lots of non-specific staining (Figure 4.17). Considering that the CER, if present, was likely to be expressed on the apical (lumen-facing) surface of the organoid cells, methods allowing the bait protein to access this surface were pursued hereafter.

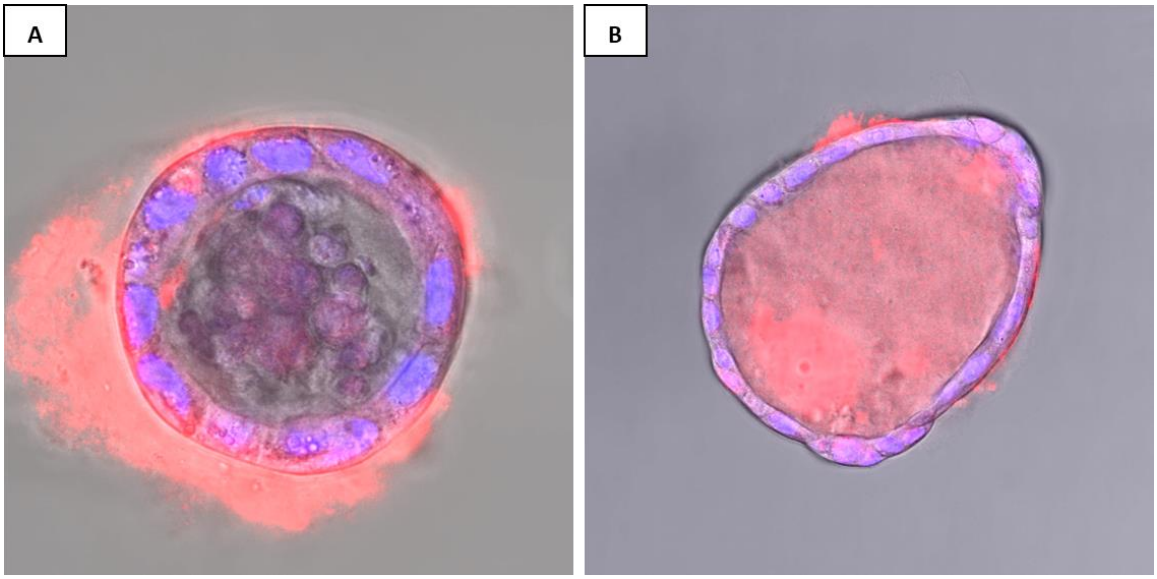

Figure 4.17. Examination of feline intestinal organoids by immunofluorescence assay by bait protein recognition. Feline intestinal organoids were cultured in chamber slides, fixed *in situ* and incubated with IgG Fc – negative control (A) or the Type 1 bait protein (B). The organoids were then incubated with a CF 633-conjugated secondary antibody (red). Nuclei were stained with DAPI (blue). Cells were then viewed using a widefield fluorescence and phase contrast microscope with a 40x oil immersion objective.

A method was trialled whereby the organoids were broken up mechanically by pipetting them and passing them through a needle before they were applied to glass slides by cytocentrifugation. Two variations of this method were tested: where the broken-up organoids were incubated with bait protein before being fixed ('incubate first' protocol), and where they were incubated with bait protein after ('fix first' protocol; Figure 4.18). Cell surface staining with the Type 2 bait protein was observed in a small population of cells under both protocols, though staining was brighter when cells were fixed first. Bright, stippled cytoplasmic staining with the Type 1 bait protein was observed in a small population of cells only when they were incubated with bait protein before being fixed, therefore organoids from different cats and passage numbers were screened using the 'incubate first' protocol. This experiment was repeated 12 times, and typical results are shown in Figure 4.19.

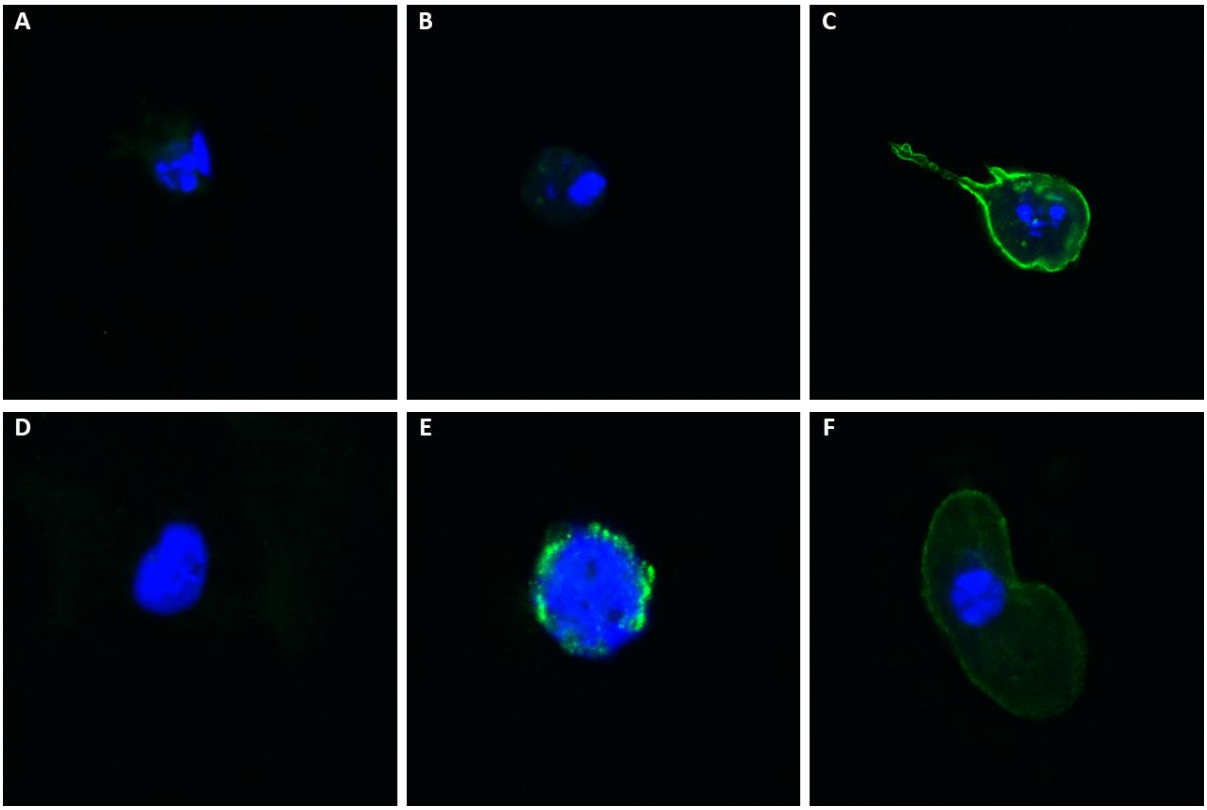

Figure 4.18. Comparing two protocols for examination of single cells from feline intestinal organoids by immunofluorescence assay for bait protein recognition. Feline intestinal organoids were mechanically dissociated. For conditions A-C, the cells were applied to glass slides by cytocentrifugation, fixed without permeabilisation then incubated with IgG Fc – negative control (A), Type 1 bait protein (B) or Type 2 bait protein (C; 'fix first' protocol). For conditions D-F, the cells were incubated with IgG Fc – negative control (D), Type 1 bait protein (E) or Type 2 bait protein (F) then applied to glass slides by cytocentrifugation and fixed without permeabilisation ('incubate first' protocol). In all conditions the cells were then incubated with a DyLight 488-conjugated secondary antibody (green). Nuclei were stained with DAPI (blue). Cells were then viewed using a confocal laser scanning microscope with a 40x oil immersion objective.

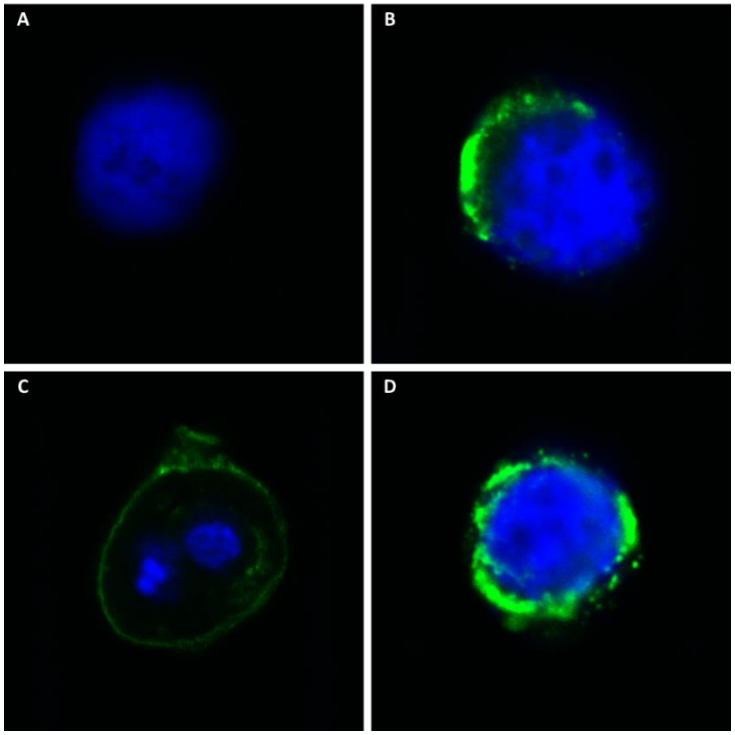

Figure 4.19. Examination of single cells from feline intestinal organoids by immunofluorescence assay for bait protein recognition. Feline intestinal organoids were mechanically dissociated. The cells were incubated with IgG Fc – negative control (A), Type 1 bait protein (B) or Type 2 bait protein (C+D) then applied to glass slides by cytocentrifugation and fixed. The cells were then incubated with a DyLight 488-conjugated secondary antibody (green). Nuclei were stained with DAPI (blue). Cells were then viewed using a confocal laser scanning microscope with a 40x oil immersion objective. These results are typical of 12 replicates.

Across organoids from different cats and passage numbers, the following staining pattern was consistently observed: a very small population of cells showed bright, stippled cytoplasmic staining with the Type 1 (B) and 2 (D) bait proteins, and a distinct small population of cells showed fainter surface staining with the Type 2 bait protein only (C). Cells staining with the T1 bait protein appeared to have a rounded morphology with little cytoplasm, but since a cell surface marker was not used it could not be ruled out that the staining was just perinuclear and the cell was larger than it appeared to be. To address this question, an image of a positive cell was taken under a light microscope to get a better idea of its morphology (Figure 4.20). This confirmed that the staining was cytoplasmic and the cell has very little cytoplasm, rather than the staining just being perinuclear.

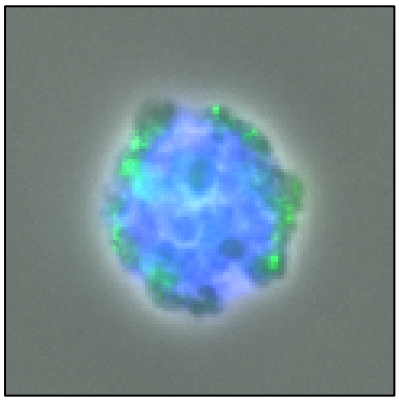

Figure 4.20. Image of a single feline intestinal organoid cell that is positive for staining with the Type 1 bait protein, taken to assess the morphology of the cell. Feline intestinal organoids were mechanically dissociated. The cells were incubated with the Type 1 bait protein then applied to glass slides by cytocentrifugation and fixed. The cells were then incubated with a DyLight 488-conjugated secondary antibody (green). Nuclei were stained with DAPI (blue). Cells were then viewed using a widefield fluorescence and phase contrast microscope with a 40x oil immersion objective.

4.2.7 Proteomic analysis of the proteins precipitated from organoids by bait proteins

If the Type 1 bait protein was recognising a CER on feline intestinal organoid cells, the staining would be expected to be cell surface rather than cytoplasmic. However, the organoid cells had been processed in such a way that could affect their morphology, and it is possible that a Type 1 FCoV CER is only expressed at the cell surface transiently. Since there seemed to be a consistent interaction between the Type 1 bait protein and a component of the organoid cells, identification of that component was attempted using immunoprecipitation followed by mass spectrometry analysis (as optimised in section 4.2.2).

Lysates were prepared from feline intestinal organoids from different cats and at different passage numbers, and these lysates were incubated with the Type 1 and 2 bait proteins, and, as a negative control, human IgG Fc. The bait proteins and IgG Fc in turn were then captured with Protein A resin. Proteins eluted from the resin were then separated by SDS-PAGE prior to in-gel digestion, TMT-labelling, mass spectrometry and proteomic analysis. Across seven replicates of this experiment comprising different cat and passage numbers, the mean number of proteins identified was 118 (±15.1; Appendix D). Eight proteins were significantly enriched (i.e. paired sample t-test p value <0.05) across all replicates with the Type 1 bait protein (Table 4.4), whereas only four were significantly enriched with the Type 2 bait protein (Table 4.5).

Table 4.4. The proteins significantly enriched by the Type 1 bait protein from feline intestinal organoids across all replicates. A consensus is shown from seven replicates. Proteins were identified following in-gel digestion, TMT-labelling and mass spectrometry analysis of the immunoprecipitates, then sorted based on the log2 fold change in their abundance compared with the negative control immunoprecipitation. Only proteins that were significantly elevated across all replicates ($p \leq 0.05$) are shown.

Uniprot accession number	Protein description	Log fold change over control	Paired sample t-test *p* value
M3W0U2	Putative serine/threonine-protein kinase Nek3	5.80	<0.001
A0A337RXE8	Putative heat shock 70 kDa protein 1A	3.35	0.003
M3W6I2	Putative actin, cytoplasmic 2	1.98	0.005
M3VZ17	40S ribosomal protein SA	1.87	0.017
A0A2I2UZM2	Putative heat shock cognate 71 kDa protein	1.79	0.007
A0A2I2UTE6	Putative alpha-enolase	1.64	0.009
Q66RN5	Elongation factor 1-alpha 1	1.49	0.034
M3W7A1	Putative actin, aortic smooth muscle	1.49	0.038

Table 4.5. The proteins significantly enriched by the Type 2 bait protein from feline intestinal organoids across all replicates. A consensus is shown from seven biological replicates. Proteins were identified following in-gel digestion, TMT-labelling and mass spectrometry analysis of the immunoprecipitates, then sorted based on the log2 fold change in their abundance compared with the negative control immunoprecipitation. Only proteins that were significantly elevated across all replicates ($p \leq 0.05$) are shown.

Uniprot accession number	Protein description	Log fold change over control	Paired sample t-test *p* value
M3VW27	Putative 40S ribosomal protein S25	1.28	0.005
Q66RN5	Elongation factor 1-alpha 1	1.15	0.009
M3W6I2	Putative actin, cytoplasmic 2	0.93	0.008
A0A2I2UTE6	Putative alpha-enolase	0.67	0.005

4.2.8 Validation of HSPA1A as a CER for Type 1 FCoV

Of the proteins enriched by the Type 1 bait protein from feline intestinal organoids, the strongest interactor was putative serine/threonine-protein kinase Nek3 (Nek3). However, putative heat shock 70 kDa protein 1a (HSPA1A) was pursued as a potential Type 1 FCoV CER because it appeared to be the most biologically relevant, having been identified as a putative CER or attachment factor for other viruses (Reyes-Del Valle *et al.*, 2005, Zhu *et al.*, 2012, Das *et al.*, 2009, Pujhari *et al.*, 2019, Khachatoorian *et al.*, 2018) and part of the receptor complex for the Gammacoronavirus IBV (Zhang *et al.*, 2017). In contrast, Nek3 has not been identified as a CER and is not known to be present on the plasma membrane. Putative heat shock cognate 71 kDa protein, also known as Hsc70, was the fifth most enriched protein from feline intestinal organoids. Hsc70 is considered to be the constitutively expressed form of HSPA1A and has been identified as a receptor for rotavirus (Perez-Vargas *et al.*, 2006).

A small proportion of each immunoprecipitate was not submitted for mass spectrometry and proteomic analysis, but instead was analysed by SDS-PAGE and western blot for the presence of HSPA1A. However, a band of the correct size was not found and the results were inconclusive (data not shown). Since there is not an anti-HSPA1A antibody validated for use in cats, it was assumed that the antibody used in the western blot may not have either recognised the protein of interest or allowed sufficient sensitivity of detection by western blot analysis.

Next, CrFK cells were heat shocked to induce expression of HSPA1A (Bilog *et al.*, 2019). An immunofluorescence experiment was carried out on the heat shocked cells to determine whether they would show surface staining with the Type 1 bait protein compared to non-heat shocked cells (Figure 4.21). Neither heat shocked nor non-heat shocked cells were stained by the Type 1 bait protein, suggesting that heat shock did not induce expression of a protein that could be detected by the Type 1 bait protein. Again, since an antibody known to be capable of recognising feline HSPA1A is not available, it is unknown whether the heat shocked cells failed to express HSPA1A or if HSPA1A was expressed but not recognised by the Type 1 bait protein.

To explore the interaction between the Type 1 bait protein and HSPA1A, a different approach was taken, whereby recombinant HSPA1A was expressed in CrFK cells. A plasmid containing the feline HSPA1A coding sequence was synthesised and the coding sequence was then subcloned into mammalian expression vector pcDV4 to create a plasmid termed pHSPA1A. To circumvent the problem of not having an anti-HSPA1A antibody validated for use in cats, sequence encoding a FLAG (polypeptide) tag was fused to the 3´ HSPA1A sequence, allowing the recombinant protein to be detected using a commercially available anti-FLAG antibody. See Appendix A for the sequence of

pHSPA1A. In order to determine the optimal transfection conditions, CrFK cells were transfected with pHSPA1A using a range of Lipofectamine concentrations, fixing the cells at two different time points and carrying out an immunofluorescence assay with and without permeabilisation (Figure 4.22). The strongest expression of HSPA1A appeared to be in cells that were transfected with a high Lipofectamine concentration, fixed at 24 hours and permeabilised. Although the expression of a Type 1 FCoV CER would be expected to be cell surface, expression of recombinant HSPA1A in transfected cells was mostly cytoplasmic so permeabilisation to allow the bait protein to reach HSPA1A seemed more important than preserving the plasma membrane.

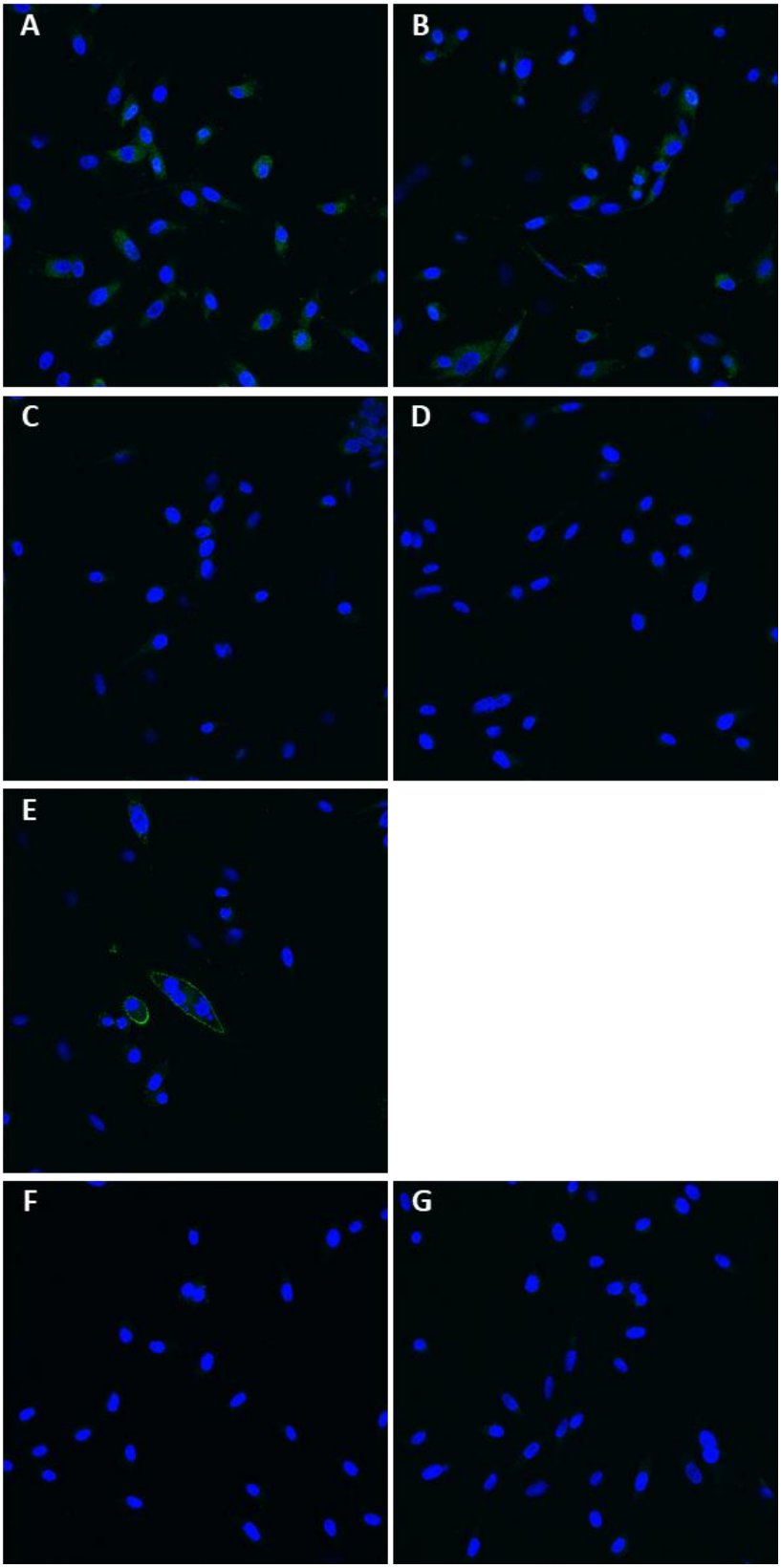

Figure 4.21. Examination of heat shocked CrFK cells by immunofluorescence assay for bait protein recognition. The CrFK cells were grown on glass coverslips and either cultured under normal conditions (A, C, E, F) or heat shocked at 42°C for two hours then recovered at 37°C for 16 hours (B, D, G). Cells were fixed without permeabilisation then incubated with IgG Fc – negative control (A, B), Type 1 bait protein (C, D), Type 2 bait protein (E) or an anti-HSPA1A antibody (F, G). This was followed by a DyLight 488-conjugated secondary antibody (green) and nuclei were stained with DAPI (blue). Cells were then viewed using a confocal laser scanning microscope with a 40x oil immersion objective.

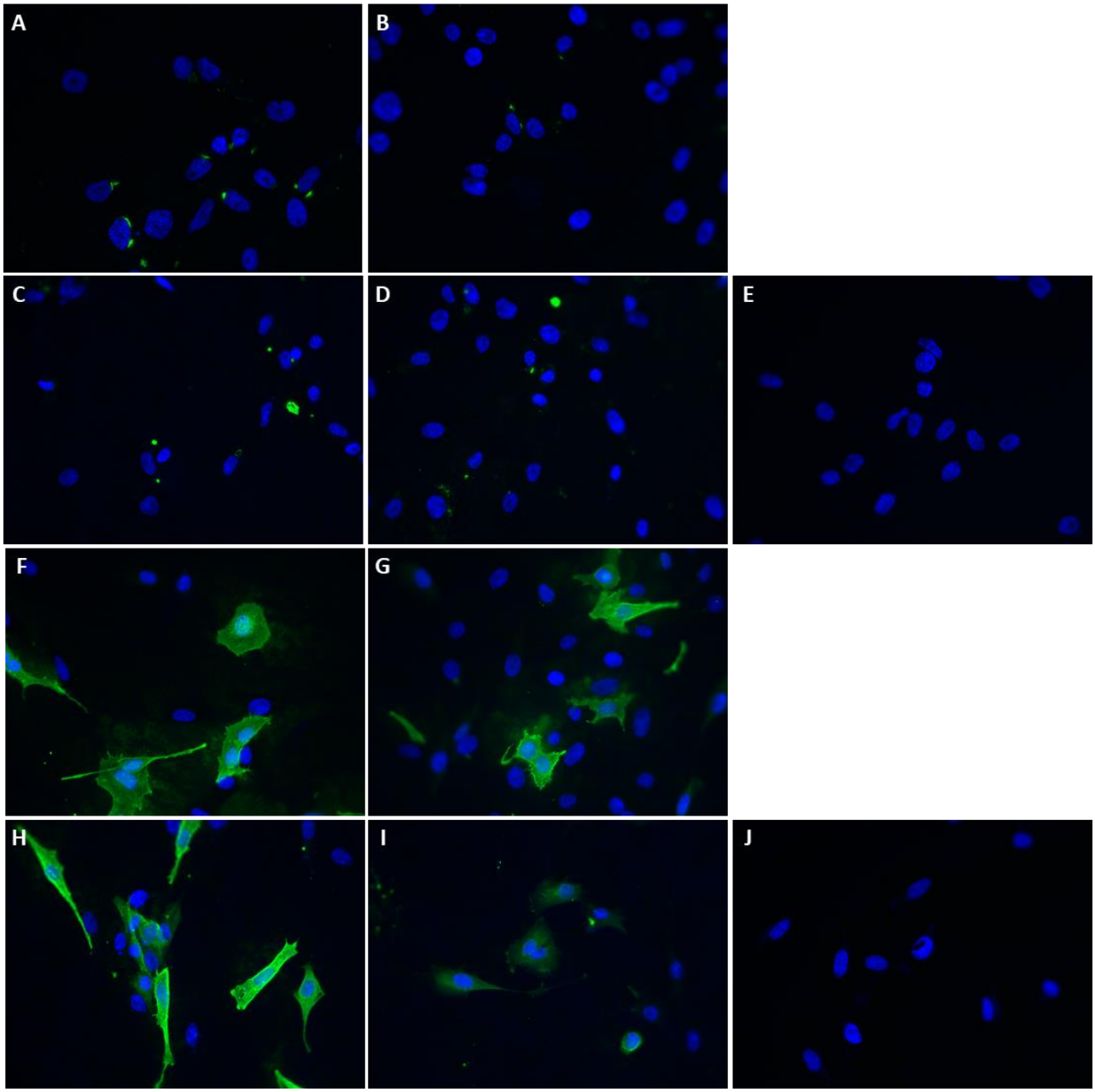

Figure 4.22. Examination of pHSPA1A-transfected CrFK cells by immunofluorescence assay for HSPA1A-FLAG recognition. The CrFK cells were grown on glass coverslips, transfected with pHSPA1A using a low (A-B, F-G) and high (C-E, H-J) concentration of Lipofectamine, then fixed at 24 hours (A, C, E, F, H, J) or 48 hours (B, D, G, I) with (F-J) and without (A-E) permeabilisation. The cells were incubated with an anti-FLAG antibody (A-D, F-I) or blocking buffer only as a negative control (E, J), followed by a DyLight 488-conjugated secondary antibody (green). Nuclei were stained with DAPI (blue). Cells were then viewed using a widefield fluorescence microscope with a 40x oil immersion objective.

A: Low lipofectamine concentration, fixed at 24 hours, not permeabilised, anti-FLAG antibody;
B: Low lipofectamine concentration, fixed at 48 hours, not permeabilised, anti-FLAG antibody;
C: High lipofectamine concentration, fixed at 24 hours, not permeabilised, anti-FLAG antibody;
D: High lipofectamine concentration, fixed at 48 hours, not permeabilised, anti-FLAG antibody;
E: High lipofectamine concentration, fixed at 24 hours, not permeabilised, negative control;
F: Low lipofectamine concentration, fixed at 24 hours, permeabilised, anti-FLAG antibody;
G: Low lipofectamine concentration, fixed at 48 hours, permeabilised, anti-FLAG antibody;
H: High lipofectamine concentration, fixed at 24 hours, permeabilised, anti-FLAG antibody;
I: High lipofectamine concentration, fixed at 48 hours, permeabilised, anti-FLAG antibody;
J: High lipofectamine concentration, fixed at 24 hours, permeabilised, negative control.

On establishing the optimal conditions for CrFK transfection with pHSPA1A, an immunofluorescence assay was carried out using these conditions to determine whether the Type 1 bait protein recognised HSPA1A expressed in CrFK cells (Figure 4.23). The assay demonstrated no interaction between pHSPA1A-transfected cells and Type 1 bait protein.

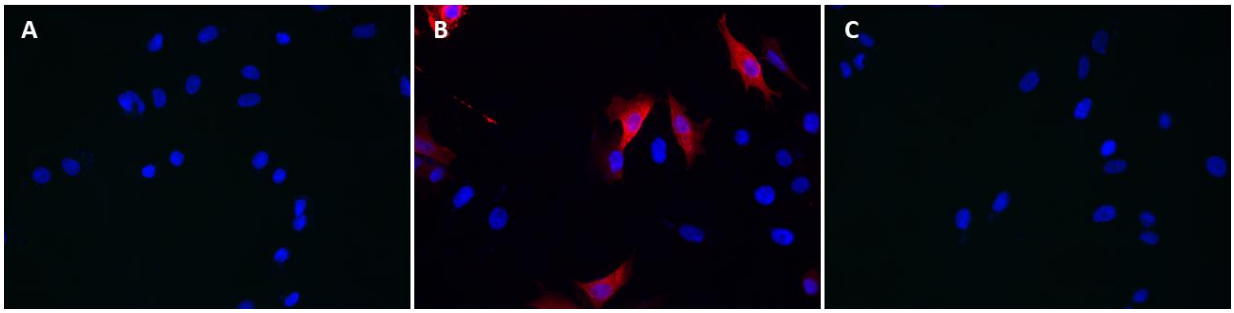

Figure 4.23. Examination of pHSPA1A-transfected CrFK cells by immunofluorescence assay for Type 1 bait protein recognition. The CrFK cells were grown on glass coverslips, transfected with pHSPA1A then fixed at 24 hours with permeabilisation. The cells were incubated with IgG Fc – negative control (A), an anti-FLAG antibody (B) or Type 1 bait protein (C), followed by a DyLight 488-conjugated anti-Fc secondary antibody (green) and a CF 633-conjugated anti-FLAG secondary antibody (red). Nuclei were stained with DAPI (blue). Cells were then viewed using a widefield fluorescence microscope with a 40x oil immersion objective.

Finally, an immunoprecipitation assay was carried out whereby pHSPA1A-transfected CrFK cells were lysed and incubated with the Type 1 and 2 bait proteins, and, as a negative control, human IgG Fc. These proteins in turn were incubated with Protein A resin, and the proteins eluted from the resin were analysed by SDS-PAGE and western blot (Figure 4.24). This demonstrated a protein of around 70 kDa (identified by the FLAG tag) present only in the transfected cells, which was assumed to be HSPA1A. This protein was extremely abundant in the transfected cell lysate, and less abundant but still present in all transfected cell immunoprecipitates. When normalised by the intensity of the bait protein (or human IgG Fc for negative control) band in that lane, the quantified intensity of the HSPA1A band was 0.56 with Type 1 bait protein, 0.26 with Type 2 bait protein and 0.16 with negative control. This demonstrates a 3.5-fold enrichment of HSPA1A with the Type 1 bait protein compared to the negative control.

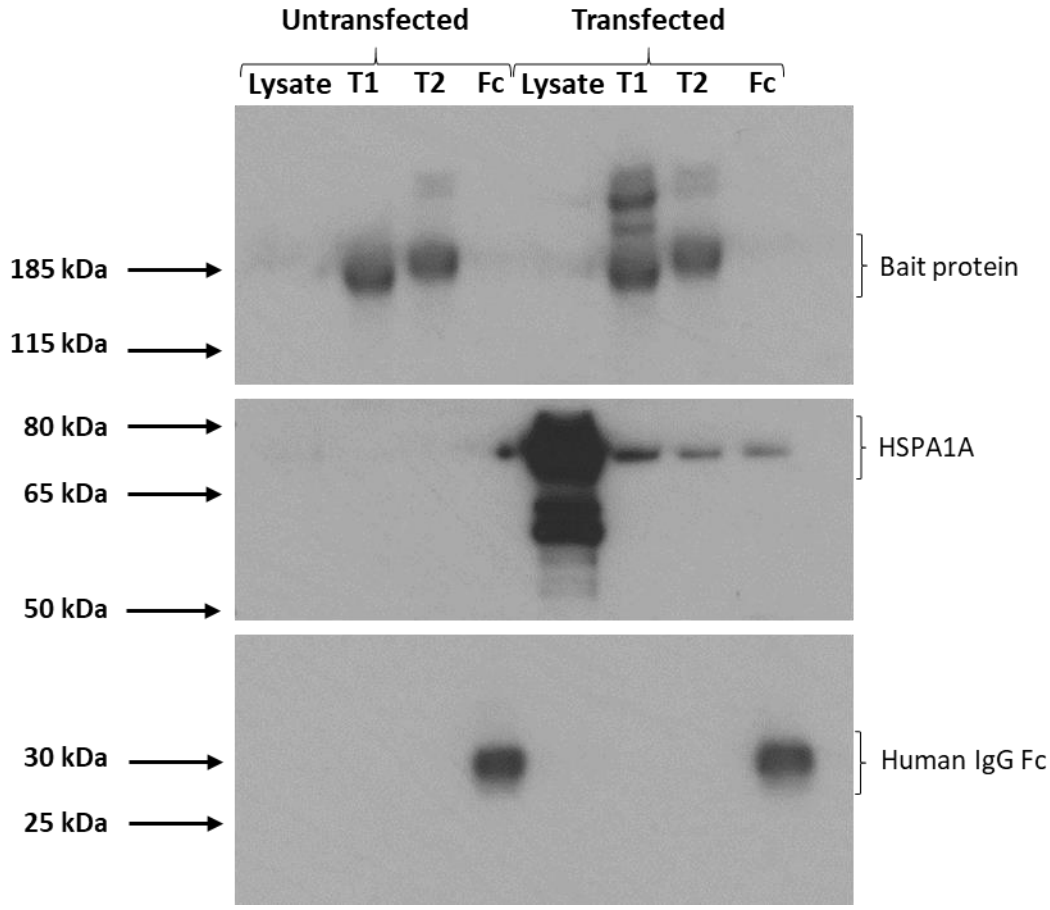

Figure 4.24. Analysis by western blot of the proteins immunoprecipitated from pHSPA1A-transfected cells by the bait proteins. Lysates of pHSPA1A-transfected and untransfected CrFK cells were divided into three and incubated with Type 1 bait protein (T1), Type 2 bait protein (T2) or IgG Fc – negative control (IgG). These in turn were incubated with Protein A resin, which binds the IgG part of the bait proteins or IgG Fc. Proteins at different stages of the procedure were denatured and run on a 4-12% SDS-PAGE Bis-Tris mini gel. Proteins on the gel were transferred on to a membrane, which was probed with anti-human IgG Fc antibody (upper and lower sections of the membrane) or anti-FLAG antibody (middle section of the membrane). The membrane was then incubated with a horseradish peroxidase-labelled secondary antibody, prior to application of a chemiluminescent substrate and exposure to x-ray film. The positions of relevant molecular mass markers are shown in kDa.

4.3 Discussion

This chapter describes an investigation aiming to identify a Type 1 FCoV CER. Knowledge of the receptor utilised by Type 1 FCoV would constitute a step towards *in vitro* propagation of strains of FCoV representative of natural infections. The production of a cell-based system to grow such strains *in vitro* would facilitate our understanding of the virus, as well as enabling rescue of recombinant viruses as part of a reverse genetic system.

The 'bait protein' method – developed by Raj *et al.* (2013) to identify a MERS-CoV receptor – was utilised in the quest to identify a Type 1 FCoV CER. Raj *et al.* (2013) incubated MERS-CoV bait protein with lysates of Vero and Huh7 cells, which are both susceptible to infection with the virus. This approach was not possible in this project, because there are no cell lines available to us that are susceptible to infection with Type 1 FCoV. However, there are many available cell lines (including CrFK) that carry fAPN and are susceptible to infection with Type 2 FCoV. Both Type 1 and Type 2 bait proteins were therefore produced in this project, and the method was validated and optimised using the Type 2 bait protein with CrFK cells. The 'identification' of fAPN in this project was encouraging, as it demonstrated that the Type 2 bait protein was produced in a conformationally faithful, active form and suggested that the Type 1 bait protein might also be capable of recognising its receptor. An assumption made in this project is that a CER for Type 1 FCoV would be a protein as opposed to a carbohydrate. The known CERs for all Alphacoronaviruses and most coronaviruses in general are proteins (Jaimes and Whittaker, 2018), but carbohydrates have been reported as receptors for some Beta- and Gammacoronaviruses (Li, 2015). However, current evidence suggests that in some cases these molecules serve as attachment factors rather than CERs (Szczepanski *et al.*, 2019), allowing the virus to attach to the host cell but not mediating cell entry. If Type 1 FCoV used a carbohydrate as its CER this molecule may not be recognised by our approach, but on the balance of evidence this situation seems very unlikely.

The next step towards identifying a Type 1 FCoV CER was to find a cell type expressing the receptor. Cell lines that have been identified as resistant to infection with Type 1 FCoV include FCWF, feline kidney Colorado University, HEK293T and swine testicular cells (Dye *et al.*, 2007), so these were not screened. Two cell lines that had not been previously investigated – DH82 and FE-A – were, however, screened, which identified specific staining of DH82 cells with the Type 2 but not Type 1 bait protein. As cell lines carrying the Type 2 FCoV receptor are already available, the interaction between DH82 cells and Type 2 bait protein was not pursued. Developed originally to isolate and propagate feline leukaemia virus (Jarrett *et al.*, 1973), FE-A cells are feline whole embryo cells. They are currently used by some laboratories to culture Type 2 FCoV, for the purpose of measuring FCoV

antibody serum titre by immunofluorescent antibody test (Addie *et al.*, 2015). In this study, two separate experiments demonstrated no staining of FE-A cells with either Type 1 or Type 2 bait protein. These two ideas are at odds, because in order to be permissive to Type 2 FCoV, this cell type must carry a receptor recognised by the Type 2 FCoV spike protein. It is unknown why FE-A cells were not recognised by Type 2 bait protein, but some possible explanations could include FE-A cells only expressing APN under certain conditions, FE-A cells expressing APN below the limit of detection for the assay or cells other than FE-A being mistakenly used in the assay.

Monocytes were examined next as a cell type potentially expressing a Type 1 FCoV CER. As well as being infected by the virus in the natural course of infection (Kipar and Meli, 2014), feline monocytes are permissive to Type 1 FCoV *in vitro* (Van Hamme *et al.*, 2007), albeit to a strain – 'Black' – that is so adapted to cell culture it can no longer be said to represent Type 1 FCoVs found in the field (Pedersen, 2009). Additionally, preliminary work by Dye (2006) suggested that viral pseudotypes expressing a Type 1 FCoV field strain S protein were able to infect *ex vivo* feline monocytes. After encountering issues with non-specific staining when using feline PBMC, live monocytes were selected by adherence to a culture surface and allowed to mature. As well as selecting for monocytes, this approach excluded the dead cells and debris that may have been causing false positive results with PBMC. A preliminary immunofluorescence experiment suggested that both the Type 1 and 2 bait proteins may be specifically binding to the surface of a small population of these monocyte-derived cells, but this was not consistently repeatable across experiments and was at times difficult to differentiate from the negative control. It would have been interesting to pursue this further, but the PBMC ran out and the results were not promising enough to warrant collecting more.

Feline intestinal epithelial cells were explored next, as it is known that FCoV enters these cells during the natural course of infection (Kipar *et al.*, 1998b). The S1 portion of the Type 1 bait protein was based on FCoV strain 80F, which was isolated from the faeces of a healthy, FCoV-infected cat (Lewis *et al.*, 2015). Though Type 1 FCoV is known to infect monocytes, using intestinal epithelial cells may be more relevant because a virus isolated from the faeces would be expected to have a very strong tropism for the intestinal epithelium. Desmarets *et al.* (2013) had success in isolating feline intestinal epithelial cells, culturing them and infecting them with Type 1 FCoV. Unfortunately, the intestinal epithelial cell line developed by this group was not available for this investigation. Their method of isolating and culturing intestinal epithelial cells was attempted but proved unsuccessful due to poor cell viability and mucus contamination, and it has since been learnt that the group were unable to repeat their original result (Professor H. Nauwynck, personal communication, 27 January 2017). So far there has been no report in the literature of the use of this cell line to identify a Type 1 FCoV CER.

Feline intestinal organoids – three-dimensional cell cultures containing a stem cell population that differentiates into all the cells of the intestinal epithelium – were chosen as a 'source' of feline intestinal epithelial cells. On contacting Dr M. Behnke, who recently published a report detailing the acquisition and propagation of feline intestinal organoids (Powell and Behnke, 2017), his group was extremely cooperative and willing to share feline intestinal organoid cultures. The published report describes difficulties in growing feline intestinal organoids beyond around 10 passages, but in personal correspondence Dr M. Behnke explained that this problem has since been overcome by adding a high concentration of nicotinamide to the conditioned medium. Nicotinamide is a precursor to nicotinamide adenine dinucleotide (NAD): a molecule that inhibits a class of protein deacetylases known as sirtuins, that are involved in transcriptional regulation, metabolism, apoptosis, differentiation and ageing (Denu, 2005). Treatment with nicotinamide was shown to reverse key signs of senescence in murine intestinal organoids (Uchida *et al.*, 2019) and improve the efficiency and longevity of human intestinal organoids (Sato *et al.*, 2011). Despite following the recommendation to add a high concentration of nicotinamide to the medium, the recovered feline organoid cultures in this project failed to grow robustly beyond nine passages.

At this point it was unclear where the problem lay: with one or more of the reagents used, with organoid handling technique or with the viability of the organoid cultures. When using reagents donated by Dr M. Behnke failed to make a difference, the first problem was ruled out. It seemed unlikely that the problem lay with technique, since no issues arose when cultivating murine intestinal organoids. It therefore seemed most probable that the viability of the organoid cultures was compromised, perhaps because of temperature or pressure fluctuations encountered during shipping. Organoid cultures established in-house performed better, growing robustly for 7 to 9 passages before ceasing to expand. Similarly to Powell and Behnke (2017), a mesenchymal-like cell type was observed to co-culture with the organoids until around passage 7. These cells are likely subepithelial myofibroblasts, which have been reported to support the growth, differentiation and expansion of human intestinal organoids (Lahar *et al.*, 2011). It is possible that the mesenchymal-like cells were supporting organoid growth either structurally or through secretion of signalling molecules. In contrast to mouse and human systems, feline intestinal organoid cultures are not well characterised and their growth requirements are not completely understood, so there is not a wide evidence base to draw from when establishing the correct conditions for growth.

Although it was difficult to propagate the organoids beyond several passages, there was enough opportunity while the organoids were growing robustly to screen them for a Type 1 FCoV CER and make lysates for immunoprecipitation experiments. However, if the organoids are to be used to propagate Type 1 FCoV in the future, further optimisation would be advantageous to identify

conditions for continuous growth. This could involve establishing a co-culture system with the mesenchymal-like cells, perhaps by immortalising them or renewing them every few passages. The organoids cultured in this project are technically enteroids because they are derived from intestinal stem cells and contain only an epithelial cell layer (though they are referred to in this project as organoids as this term is more commonly used and understood). In contrast, authentic organoids are derived from pluripotent stem cells and contain a mesenchymal as well as an epithelial cell layer (Ramani *et al.*, 2018). An alternative to immortalising the mesenchymal-like cells could be to develop authentic feline intestinal organoids, containing a mesenchymal layer, from feline pluripotent stem cells.

Though 3D cultures are excellent in recapitulating the organs from which the organoids are derived, the main drawback is that access to the apical (lumen-facing) surface of the cells is restricted. This is a particular problem for infection of the organoids; since intestinal pathogens access the epithelial cells *via* the intestinal lumen, their receptors would be expected to be expressed on the apical surface of the cell. This problem has been overcome through microinjection, where the pathogen is injected directly into the lumen of the organoid (Bartfeld and Clevers, 2015). The 3D nature of organoid cultures also poses a problem when attempting to access the apical surface of the cells for staining with antibodies or bait proteins. A method whereby the organoids were fixed and stained *in situ* was trialled, but, despite following the protocol used successfully by Mahe *et al.* (2013), this resulted in non-specific staining and uninterpretable results. Monolayers, organoid cultures grown in 2D rather than 3D, present a solution, because they provide access to the apical surface of the cell while still recapitulating the intestinal environment (van der Hee *et al.*, 2018), and indeed they have been used to successfully propagate intestinal viruses (Ettayebi *et al.*, 2016, Saxena *et al.*, 2016). In the future, developing feline intestinal organoid monolayer cultures would be a priority objective for furthering this project.

In the absence of monolayers, a method was developed in-house to allow access to the apical surface of the epithelial cells. In this method, organoids were liberated from the Matrigel matrix and mechanically dissociated before incubation with bait protein, application to glass slides and fixation ('incubate first' protocol). Across different cats and passage numbers, a consistently repeatable result was observed: a small population of cells showed bright, stippled cytoplasmic staining with the Type 1 and 2 bait proteins, and a distinct small population of cells showed fainter surface staining with the Type 2 bait protein only. The surface staining observed with Type 2 bait protein is the staining pattern that was expected, but the cytoplasmic staining observed with both bait proteins is a conundrum for two reasons: the morphology of this cell type is unexpected for an

epithelial cell, and the presence of staining in the cytoplasm as opposed to on the cell surface is unexpected.

The cells with cytoplasmic staining were always small, round cells with very large, fluffy nuclei and very little cytoplasm. This appearance is not typical for an epithelial cell, but perhaps more suggestive of a lymphocyte (Murphy *et al.*, 2012). Although it is feasible that lymphocytes could be carried over from blood or Peyer's patches during the crypt isolation process, this has not been reported and was not observed in our cultures. Alternatively, the cells may have been epithelial but with disturbed morphology due to being liberated from a multicellular structure or some other aspect of the immunofluorescence protocol, or they may have represented a less typical component of intestinal organoids (e.g. Paneth cells). Regardless of the cell type, cytoplasmic staining should not be present because at no point were the cells permeabilised. It is conceivable that the bait protein was taken up by live cells during incubation, but this is unlikely because the incubation step was performed at 4 °C. Besides, this still would not explain the cytoplasmic staining, because the secondary antibody was applied to fixed, non-permeabilised cells so should have not been able to get past the cell's plasma membrane. Perhaps it is more likely that the plasma membrane was compromised during the process of mechanical dissociation, allowing the bait protein to bind to an internal protein and the secondary antibody to subsequently enter the cell. To further investigate both problems, an antibody could be included in the immunofluorescence protocol to detect a cell surface epithelial antigen. If positive, this would suggest both presence of a plasma membrane and epithelial origin of the cells. Alternatively, a live/dead stain could be used to determine whether the cells have an intact plasma membrane.

A variant of the organoid immunofluorescence method was trialled, whereby the organoid cells were fixed prior to incubation with bait protein ('fix first' protocol). Interestingly, this yielded only Type 2 bait protein surface staining and not the cytoplasmic staining seen with both bait proteins. This could be because the cell's plasma membrane is not damaged by this method, the bait protein is being taken up by some other mechanism (e.g. phagocytosis) or the protein of interest is not recognisable by Type 1 bait protein when it is fixed. If it is assumed that the bait protein is recognising an intracellular target that has become accessible due to plasma membrane damage with the 'incubate first' protocol, it is conceivable that the 'fix first' protocol avoided this damage and therefore left the target inaccessible. However, this seems unlikely for two reasons. If any part of the method is going to damage the plasma membrane, it would be the shear forces generated by passing the cells through a needle, and this step is common to both protocols. Additionally, organoid cells undergoing the 'fix first' protocol that were permeabilised following fixation (data not shown) were still negative for staining with Type 1 bait protein. This result raises the question of whether

the staining with both bait proteins in the 'incubate first' method represents another mechanism such as phagocytosis rather than genuine receptor binding. An alternative mechanism of bait protein internalisation, though possible, seems unlikely, since it would require a cell type capable of phagocytosis in the organoid cultures and would not explain how secondary antibody is able to enter cells. Additionally, if phagocytosis was occurring, cells should also be positive with the IgG Fc negative control condition. The most likely option is that fixation renders the protein of interest unrecognisable to the bait protein, although this suggests that it would be worthwhile to screen DH82 cells, FE-A cells and PBMC/monocytes in an immunofluorescence experiment where they are incubated with the bait proteins prior to fixation.

Regardless of the questions raised by the nature of the staining, the Type 1 bait protein appeared to be consistently recognising a component of the organoid cells. The next step was therefore to carry out immunoprecipitation, mass spectrometry and proteomics experiments with the bait proteins and lysates prepared from feline intestinal organoids. Despite consistent surface staining of a population of cells across the immunofluorescence experiments, only four proteins were significantly enriched with the Type 2 bait protein across all seven replicates, and none of these are known virus receptors. Moreover, fAPN was not identified in any of the experiments. This may suggest that fAPN is not used as a CER by Type 2 FCoV in natural infections of intestinal epithelial cells, although this is contradicted by the fact that fAPN was found in a feline intestinal brush border preparation (Hohdatsu *et al.*, 1998). Alternatively, it may suggest a problem with the immunoprecipitation experiments. If this is the case, it seems most likely that the problem was lack of starting material. This would also explain why fewer proteins were identified proteomically with the organoids (on average 118 in each experiment) compared to CrFK cells (over 700). Due to the nature of the organoids, it was challenging to start with as many cells as would be possible when working with cell lines. This could potentially be overcome by growing the organoids on a larger scale, or pooling organoids from different cats and passage numbers.

The protocol that was optimised for CrFK cells and the Type 2 bait protein was followed when working with organoids, which involved using a detergent designed to specifically isolate cell surface proteins. In retrospect this was not the best detergent to use, as the immunofluorescence staining with the Type 1 bait protein was not on the cell surface. Despite this, eight proteins were significantly enriched using the Type 1 bait protein across all seven replicates. It was decided to focus on the 641 aa uncharacterised product of feline gene LOC105260573 (Uniprot accession number A0A337RXE8), which was inferred by homology to be feline heat shock 70 kDa protein 1a, also known as HSPA1A and Hsp70. HSPA1A is part of the HSP70 family: a group of ATP-dependent chaperones that participate in the cellular protein surveillance network and are involved in protein

folding (Mayer, 2013). HSPA1A is known to localise at the plasma membrane when cells are subjected to stress through association with the lipid phosphatidylserine (Bilog *et al.*, 2019), and its expression has been observed in a number of cell lines including Neuro-2a (mouse neuroblastoma) (Das *et al.*, 2009), chicken embryo kidney (Zhang *et al.*, 2017), U397 (human monocyte), human neuroblastoma (Reyes-Del Valle *et al.*, 2005), Huh7 and Vero cells (Zhu *et al.*, 2012). Since HSPA1A is not well characterised in cats, it is unknown whether this protein is expressed in any feline cell lines. Along with its constitutive form Hsc70, which was also enriched from feline intestinal organoids by Type 1 bait protein, HSPA1A has been identified as a putative CER or attachment factor for a range of viruses including IBV (Zhang *et al.*, 2017), rotavirus (Perez-Vargas *et al.*, 2006), Dengue virus (Reyes-Del Valle *et al.*, 2005), Japanese encephalitis virus (Zhu *et al.*, 2012, Das *et al.*, 2009) and Zika virus (Pujhari *et al.*, 2019, Khachatoorian *et al.*, 2018). This suggests that HSPA1A is a feasible candidate receptor for Type 1 FCoV.

An alternative explanation of the interaction between Type 1 bait protein and HSPA1A is that the latter is an attachment factor, rather than a CER, for Type 1 FCoV. Attachment factors are receptors that a virus can bind to in order to get a foothold at the host cell membrane, but, unlike CERs, they do not mediate viral entry into the cell (Jolly and Sattentau, 2013). Examples of attachment factors for coronaviruses include sialic acid for MERS-CoV (Widagdo *et al.*, 2019), BCoV and TGEV (Schwegmann-Wessels and Herrler, 2006), heparan sulfate for PEDV (Huan *et al.*, 2015) and DC-SIGN for Type 2 FCoV (Regan and Whittaker, 2008) and Type 1 FCoV strain Black (Van Hamme *et al.*, 2011, Regan *et al.*, 2010). It is perhaps more likely that HSPA1A is an attachment factor for Type 1 FCoV rather than a CER, because, although not in the top five proteins, HSPA1A was enriched from CrFK cells by Type 1 bait protein. This is inconsistent with HSPA1A being a CER for Type 1 FCoV because CrFK cells are not permissive to infection with this virus, although HSPA1A is known to only localise to the plasma membrane under certain conditions (Bilog *et al.*, 2019). Though identification in this project of an attachment factor rather than a CER for Type 1 FCoV would not be so much of a breakthrough, it would still be an important finding that could give insight into the tropism, pathogenesis and *in vitro* culture requirements of the virus. Interestingly, the only protein in the top five interactors with both Type 1 and 2 bait proteins in the immunoprecipitation with CrFK cells, putative SGTA (Uniprot accession number M3WN94), is a co-chaperone known to interact with HSPA1A and Hsc70. Interactions between SGTA and a range of viral proteins have also been reported, though not coronavirus S protein (Roberts *et al.*, 2015).

Despite being more enriched from feline intestinal organoids by Type 1 bait protein than HSPA1A, the 478 aa uncharacterised product of feline gene NEK3 (Uniprot accession number M3W0U2), inferred by homology to be Nek3, was not investigated further. This is because Nek3 has not been

identified as a CER, is not known to be present on the plasma membrane and was the second strongest interactor in the immunoprecipitation with CrFK cells and Type 1 bait protein. The latter point is inconsistent with Nek3 being a CER for Type 1 FCoV, since CrFK cells are resistant to this serotype. However, the interaction between Type 1 bait protein and Nek3 appears to be very strong and consistent, so in the future it would be interesting to investigate whether this protein plays a role as an attachment factor, or in the FCoV replication cycle post-internalisation. This could also be the case for the three proteins that were found to be enriched from feline organoid lysates by both Type 1 and Type 2 bait proteins: elongation factor 1-alpha, putative actin (cytoplasmic 2) and putative alpha-enolase. It is in fact more likely that the bright staining seen on immunofluorescence represents an interaction with one of these proteins, since this staining pattern was seen equally with both Type 1 and Type 2 bait proteins. Immunoprecipitation, mass spectrometry and proteomics is a much more sensitive way of identifying an interaction between two proteins than immunofluorescence, so it is possible that the interaction between Type 1 bait protein and HSPA1A in organoid cells was not detectable by immunofluorescence.

Since proteins of the HSP70 family are inducible by stress and HSPA1A is known to localise to the plasma membrane under stressful conditions (Bilog *et al.*, 2019), an attempt was made to heat shock CrFK cells to induce expression of HSPA1A. The heat shock treatment did not render the cells recognisable to Type 1 bait protein, but as an antibody known to be capable of recognising feline HSPA1A was not available it is unclear whether this represented no expression of HSPA1A or no interaction between HSPA1A and the bait protein. The results of a western blot carried out to confirm the enrichment of HSPA1A by Type 1 bait protein seen proteomically were similarly inconclusive. Only one anti-HSPA1A antibody was tested in this project: a recombinant antibody raised against human HSPA1A aa 50-300. Due to the homology of the human and feline HSPA1A sequences this antibody would be expected to recognise feline HSPA1A, but this does not consider post-translational modifications or possible splice variants that could obscure or remove this epitope. Given more time it would be useful to test other anti-HSPA1A antibodies.

That HSPA1A is induced and recruited to the plasma membrane under stressful conditions is interesting because, should HSPA1A be confirmed as a CER for Type 1 FCoV, this could go some way to explaining the different infection patterns seen in different cats. It is plausible that, for example, increased expression of HSPA1A on the surface of monocytes due to locally stressful conditions such as inflammation or co-infection could lead to increased infection of this cell type and therefore susceptibility to the development of FIP from FCoV infection. Conversely, the small proportion of cats who appear to be completely resistant to infection with FCoV (Addie and Jarrett, 2001) could have no expression of HSPA1A on their intestinal epithelial cells.

To further explore the interaction between HSPA1A and Type 1 FCoV, a FLAG-tagged feline HSPA1A expression plasmid was created and the recombinant protein expressed in CrFK cells. No interaction was found between Type 1 bait protein and the cells expressing HSPA1A on an immunofluorescence assay, but, using an immunoprecipitation assay, HSPA1A was found to be 3.5 times more abundant with the Type 1 bait protein than negative control. This demonstrates a specific interaction between the Type 1 bait protein and HSPA1A, confirming that HSPA1A is a CER or attachment factor for Type 1 FCoV. This result also supports the theory that the fixation method used for the immunofluorescence analysis rendered the protein of interest unrecognisable to the Type 1 bait protein, as there is no fixation step in the immunoprecipitation assay. Given more time, it would be informative to screen pHSPA1A-transfected CrFK cells in an immunofluorescence assay where the bait proteins are incubated with the cells prior to fixation. It is also possible that the FLAG tag could have interfered with the interaction between HSPA1A and Type 1 bait protein in the immunofluorescence assay. If a reliable anti-feline HSPA1A antibody was obtained, this possible complication could be circumvented by expressing the protein without a FLAG tag.

To take this work further, the obvious next step would be to express recombinant HSPA1A in CrFK cells and see whether this renders them infectable by field strains of Type 1 FCoV. Should this prove to be the case, HSPA1A could confidently be identified as a CER for Type 1 FCoV and work could begin to stably express this protein in CrFK cells. The resulting cell line would support the propagation of field strains of Type 1 FCoV *in vitro*, which would represent a huge leap forward in FCoV research.

5 Measuring the immunogenicity of FCoV-derived peptides using ELISpot

5.1 Introduction

This chapter will describe the optimisation and use of a feline IFN-γ ELISpot assay to measure the immunogenicity of different FCoV-derived peptides.

The ELISpot assay enables a wide range of antigens, including whole organisms, proteins and peptides, to be tested for their ability to elicit a T cell response, by the detection of cytokine secretion (Abbott *et al.*, 2012). The cells that go into the ELISpot assay are PBMC, which consist of T cells, B cells, natural killer cells and monocytes. Antigen-presenting cells in this mix (i.e. B cells, monocytes and monocyte-derived cells) display the antigen of interest to the T cells. Peptides are the simplest kind of antigen to use in ELISpot assays because they bind directly to empty major histocompatibility complex (MHC) molecules on the surface of antigen-presenting cells, whereas whole organisms and proteins must be internalised and processed before their antigens can be presented (Schmittel *et al.*, 2001).

On recognition of an antigen presented by an antigen-presenting cell, T cells become activated and release effector cytokines. One such effector cytokine is IFN-γ; IFN-γ is mainly secreted by $CD4^+$ T helper 1 (Th1) and $CD8^+$ T cells, which play a major role in the cellular immune response (Murphy *et al.*, 2012). Production of IFN-γ is therefore a marker for cellular immunity. The cytokine is captured by specific antibodies (e.g. anti-IFN-γ) that coat the ELISpot plate, and, following application of an enzyme-linked secondary antibody and a chromogen, the presence of cytokine is visualised by the appearance of a 'spot'. In an ELISpot assay, one activated T cell results in one spot, allowing activated T cells to be enumerated. Though an advantage of ELISpot is that it can detect a low frequency of antigen-specific T cells (Slota *et al.*, 2011), immunological priming to the antigen(s) used in the assay is necessary to boost the number of T cells able to recognise the antigen and increase the chance of a response. This is achieved by exposing the animal to the antigen through infection or vaccination before PBMC are taken for use in the assay.

An understanding of which parts of the FCoV proteome are most immunogenic, and the type of immune response they stimulate, will provide insight into the effect of mutations that overlap with these areas and potentially guide vaccine design. Since a robust cellular immune response appears to be important in overcoming FCoV infection (de Groot-Mijnes *et al.*, 2005, Pedersen *et al.*, 2015), work has begun to map the FCoV proteome for T cell epitopes: specific regions of antigens recognised by T cells. By engineering almost the entire genome of FCoV into recombinant vaccinia viruses and observing the ability of these recombinant viruses to stimulate T cells, de Groot-Mijnes

et al. (2005) found that the S protein of a virulent Type 2 FCoV was the main source of T cell epitopes in this virus strain.

Following a study demonstrating the ability of peptides derived from laboratory-adapted Type 1 FCoV strain 'KU-2' S2 protein to induce a Th1-type immune response in a mouse model (Satoh *et al.*, 2010), Satoh *et al.* (2011a) screened the N and S2 proteins of KU-2 and Type 2 FCoV strain '79-1146' more closely for T cell epitopes. The study exposed PBMC from cats experimentally infected with either strain to a series of overlapping peptides representing these areas, and measured IFN-γ production as an indicator of a cellular immune response. Cats were exposed to S2 peptides representing the serotype with which they had been infected. Most T cell epitopes for the Type 1 FCoV were found within the HR1 region of S2, and most T cell epitopes for the Type 2 FCoV were within the HR1 region of S2 and the IH region. Interestingly, these sites did not contain any antibody-binding epitopes, therefore would be expected to stimulate a selectively cellular immune response. Takano *et al.* (2014) screened the S1 and M proteins of the KU-2 and 79-1146 FCoV strains in the same way as Satoh *et al.* (2011a), but found only T cell epitopes that overlapped with antibody-binding epitopes. These findings suggest that the HR1 region of S2 from both Type 1 and Type 2 FCoV is a promising source of peptides for stimulating a strong, selective cellular immune response. Though ADE of disease is not thought to be an issue in cats naturally infected with FCoV (Addie *et al.*, 1995), choosing T cell epitopes that do not overlap with antibody-binding epitopes could be advantageous when designing a peptide-based vaccine against FCoV because ADE has been demonstrated in vaccinated, experimentally-infected cats (Balint *et al.*, 2014, Takano *et al.*, 2019b).

The immune response of an individual is determined by many factors. One explanation for why a strong cellular response to FCoV would be mounted in some cats and not others is that some variants of the virus contain epitopes that are better at stimulating T cells. Lewis *et al.* (2015) sequenced the genomes of six FCoVs: three from the faeces of healthy cats and three from the tissue lesions of cats with FIP. A consistent substitution of isoleucine with threonine at amino acid 1108 (I1108T), located in the HR1 region of S2, was found in all samples from FIP cats and none from healthy cats. This amino acid sits within a major T cell epitope for Type 1 FCoV (I-S2-6; KU-2 amino acid position 1089-1108) (Satoh *et al.*, 2011a). By testing peptides representing epitopes with and without this substitution for their ability to stimulate a cellular immune response, one could ascertain whether the substitution could have an influence on the immune response of the infected cat and therefore its clinical course and outcome.

The overall aim of this project is to develop a reverse genetic system for Type 1 FCoV, and one of the uses of a reverse genetic platform is the rational development of vaccines through targeted

introduction of mutations into the viral genome. Knowledge of the FCoV T cell epitopes and their variants that are best suited to stimulating a selectively cellular immune response in the host would guide rational vaccine design, thus allowing the platform to be used to its full advantage.

5.1.1 Aims

The aims of this part of the project were:

- To optimise a feline IFN-γ ELISpot assay that will screen overlapping peptides, representing the HR1 region of a field strain Type 1 FCoV S2 protein, for their ability to stimulate T cells.
- To ascertain whether the I1108T substitution alters the T cell response to the surrounding epitope by testing variant peptides with PBMC in the optimised feline IFN-γ ELISpot assay.

5.2 Results

5.2.1 Peptides synthesised

Twenty-eight overlapping 15-mer peptides ('HR1 peptides') were synthesised to span the HR1 region of the Type 1 FCoV strain '80F' S2 protein (Appendix C). This strain was identified in the faeces of a healthy, FCoV-infected cat (Lewis *et al.*, 2015). Peptides of 15 amino acids were chosen because they should be able to bind to both MHC-I and MHC-II, stimulating $CD8^+$ and $CD4^+$ T cells respectively. Overlapping peptides ensured that no epitopes were missed (Abbott *et al.*, 2012).

Peptides 28-33 covered the I1108T substitution that sits within this region, with peptides 28, 30 and 33 representing the isoleucine variant ('isoleucine peptides') and peptides 29, 31 and 32 representing the threonine variant ('threonine peptides'; Table 5.1).

Table 5.1. Sequences of the six overlapping peptides spanning the site of the isoleucine to threonine mutation. Peptides 28, 30, 33 represent the isoleucine variant, peptides 29, 31, 32 represent the threonine variant.

Peptide	Sequence	Variant
28	H-LALGKVSNAITTISD-OH	Isoleucine
29	H-LALGKVSNAITTTSD-OH	Threonine
30	H-VSNAITTISDGFNTM-OH	Isoleucine
31	H-VSNAITTTSDGFNTM-OH	Threonine
32	H-AITTTSDGFNTMTTI-OH	Threonine
33	H-AITTISDGFNTMTTI-OH	Isoleucine

5.2.2 Obtaining PBMC

Surplus blood was available from cats who had blood samples submitted for FCoV antibody titre testing free of charge. The free testing was initially run through a local veterinary practice and the Langford Vets Small Animal Referral Hospital, University of Bristol, where the veterinarian caring for the cat gained consent from the cat's guardian, took approximately 1 ml of blood from the cat into an EDTA tube and submitted the sample to us. When the samples were received, they were processed within two days into PBMC and plasma.

The entire yield of viable PBMC from each sample was used fresh in the ELISpot assay; freezing was trialled but was found to severely impact the recovery of viable cells. The plasma from each sample was sent to Veterinary Diagnostic Services, University of Glasgow, and underwent an indirect immunofluorescent antibody test to measure FCoV antibody titre, using Type 2 FCoV grown in FE-A cells as the antigen. This test is reported to have 100% sensitivity and specificity (Addie *et al.*, 2015) but does not distinguish between Type 1 and 2 FCoV. The FCoV antibody titre obtained for each cat was relayed to the veterinarian who originally submitted the sample. By offering FCoV serology, cats who had been immunologically primed to FCoV (i.e. were seropositive) could be identified and taken forward into the ELISpot assay.

Though this system of sample recruitment worked well, recruitment of cats was extremely slow. A new strategy was devised whereby cat breeders were offered the FCoV antibody titre test for any of their cats free of charge. Since FCoV is often present, and can be an issue, in multi-cat environments such as breeding households (Drechsler *et al.*, 2011), we correctly predicted that breeders would be very interested in taking up the testing. Information about the project and the laboratory's contact details were disseminated *via* a newsletter to cat breeders, which resulted in a very large number of responses. Interested breeders were advised to make an appointment with their veterinarian, and blood sampling, submission and communication of results were coordinated through the veterinarian. Most cats recruited in this way were from the UK to enable samples to be processed as quickly as possible, but one breeder resided in Norway and another in Italy.

5.2.2.1 Population description

In total 110 blood samples were received from 110 cats before recruitment was terminated due to time limitations. Twenty-seven cats (25%) were from non-breeding households ('non-breeder') and 83 (75%) were from breeding households ('breeder'). A plasma antibody titre was obtained for all 110 cats to determine their exposure to FCoV. Thirty-two (29%) had a titre of 0 (i.e. they were seronegative), three (3%) had a titre of 80, seven (6%) had a titre of 160, four (4%) had a titre of 320, eight (7%) had a titre of 640, 31 (28%) had a titre of 1280 and 25 (23%) had a titre of >1280 (Figure

5.1). Just two (7%) non-breeder cats were seropositive, compared to 76 (92%) breeder cats (p<0.001).

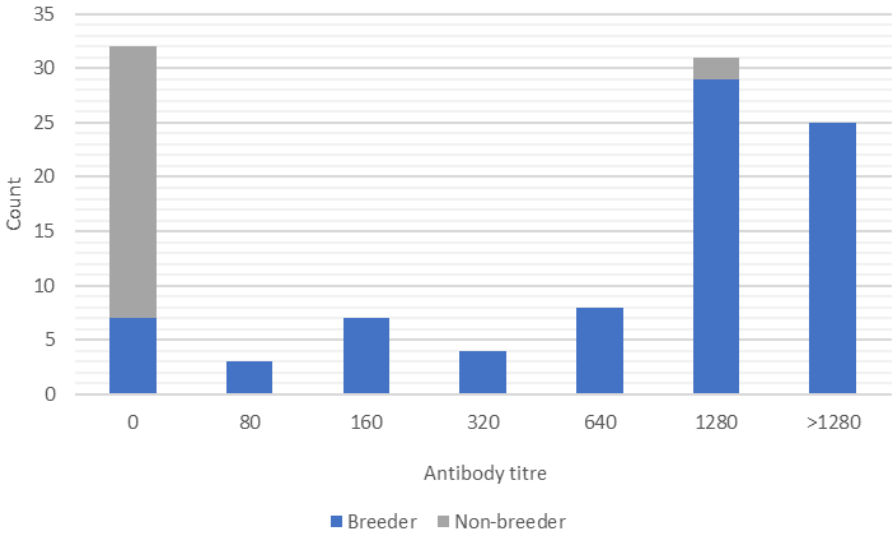

Figure 5.1. The FCoV antibody titres and source of the individual cats whose blood samples were collected for this study. Whole blood from 110 cats was collected into EDTA, and the plasma was removed after allowing the blood cells to settle by gravity. The plasma was submitted to an external laboratory for immunofluorescent antibody testing. The number of cats with each plasma FCoV antibody titre is shown as a histogram plot. The numbers of cats from 'Breeder' or 'Non-Breeder' households are indicated by the shading.

The PBMC of 82 cats were used in the ELISpot assay, with 26 (both FCoV seronegative and seropositive) of these used for optimisation purposes and the remaining 56 (FCoV seropositive only) used for experimental purposes. Samples from FCoV seronegative cats were not used for experimental purposes because they may have been naïve to FCoV and so sufficient T cells to respond to FCoV antigens would not be expected to be present. Otherwise, the 56 cats were not selected for any particular reason; samples were simply processed and used in the order that they arrived. Of the 56 cats, five could not be included in the analysis due to a failure of the positive control (section 5.2.3). Of the remaining 51 cats, six were 0-6 months, 26 were 7 months-2 years, 17 were 3-6 years, one was 9 years and one was 13 years of age. Eighty-eight percent were unneutered, likely reflecting the source of the cats as all but one were used as breeding cats (Figure 5.2). Twenty-one cats were British Shorthair, 10 were Ragdoll, five were Maine Coon, three were Birman, three were Persian and nine were 'other' pedigrees (British Longhair, Domestic Longhair, Burmese, American Curl, Bengal or Egyptian Mau; Figure 5.3).

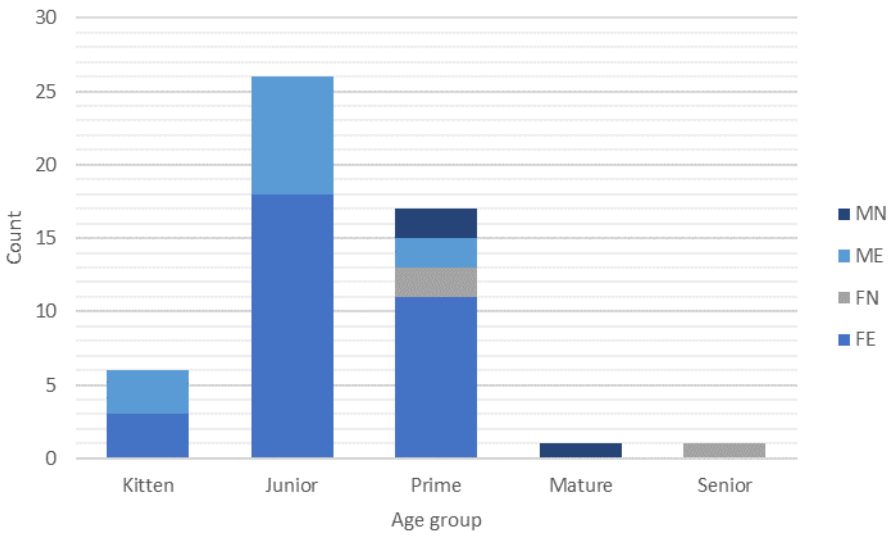

Figure 5.2. The age groups and sexes of the 51 cats whose PBMC were used in the ELISpot study. The number of cats within each age group is shown as a histogram plot. The number of cats of each sex are indicated by the shading. MN: male neutered, ME: male entire, FN: female neutered, FE: female entire. Ages are categorised as per the International Cat Care 'Life Stages' guidelines (ICC, 2015); Kitten: 0-6 months, Junior: 7 months-2 years, Prime: 3-6 years, Mature: 7-10 years, Senior: 11-14 years.

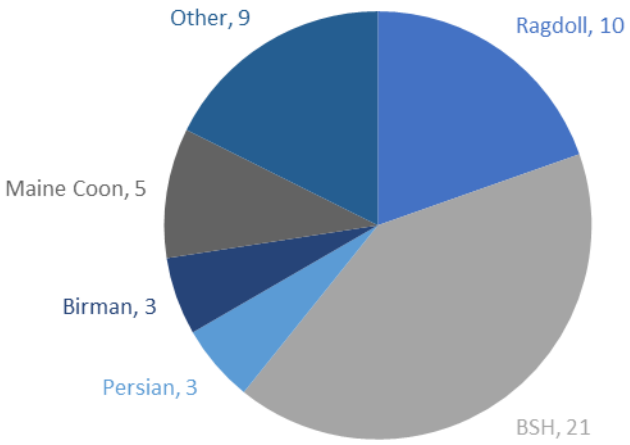

Figure 5.3. The breeds of the 51 cats whose PBMC were used in the ELISpot study. Breeds for which there were less than three cats are grouped under 'other' and are British Longhair, Domestic Longhair, Burmese, American Curl, Bengal and Egyptian Mau. BSH: British Shorthair.

5.2.3 Optimisation of ELISpot assay

The feline IFN-γ ELISpot assay (Figure 5.4) was carried out following the protocol provided by the kit manufacturer.

Figure 5.4. Schematic to illustrate a feline IFN-γ ELISpot assay. Each step depicts one well of a 96 well plate.
1: Wells of a plate are coated in anti-IFN-γ antibody (capture antibody; orange). Feline PBMC (blue) are added to the well, along with the peptide(s) of interest. Antigen-presenting cells present the peptide(s) to T cells, but this aspect of the assay is not represented here.
2: If T cells recognise the peptide(s), they become activated (shown as enlarged and green) and secrete IFN-γ (green rhombi). The IFN-γ binds to the anti-IFN-γ antibody in the vicinity of the activated cell.
3: The well is washed to remove the cells and any unbound IFN-γ. An enzyme-linked secondary anti-IFN-γ antibody (detection antibody; orange) is applied, and this binds to any IFN-γ that was captured by the capture antibody. A substrate is applied that reacts with the detection antibody, resulting in a purple spot. One spot is formed per activated T cell.

To optimise the assay before using it for experimental purposes, serial dilutions of feline PBMC, all stimulated with the mitogen phytohaemagglutinin (PHA) at 10 µg/ml, were used. This dose of PHA is expected to stimulate T cells to produce IFN-γ in the absence of antigen (Passlack *et al.*, 2017). It was found that spots were visible with 10,000 cells per well and above, but that it was difficult to distinguish spots from background colouration. To reduce the level of background colouration, an ethanol pre-wetting step was trialled based on the plate manufacturer's advice, but pre-wetted wells did not show any spot formation. Various other steps were therefore taken to reduce background colouration and/or increase the intensity of the spots, as follows:

- Incubation lengths of PBMC with PHA from 18 to 48 hours were tested, based on a range recommended by Nordone *et al.* (2005).
- The PBMC were rested for 24 hours prior to use in the assay, based on advice from Nordone *et al.* (2005).
- The underdrain was removed and both sides of the plate rinsed under tap water after the final incubation period with the chromogen, based on advice from Nordone *et al.* (2005).
- Colour development reagents were brought to room temperature before use, based on the manufacturer's recommendation (R&D, 2016).
- The plate was dried for at least 24 hours at 2-8 °C, based on advice from Abcam (2016).

Upon the testing of these conditions, incubating PBMC for 24 hours and not resting them prior to use in the assay were found to result in the most intense spots (Figure 5.5). Rinsing the plate, bringing the colour development reagents to room temperature and drying the plate were found to reduce background colouration, so were incorporated into the protocol.

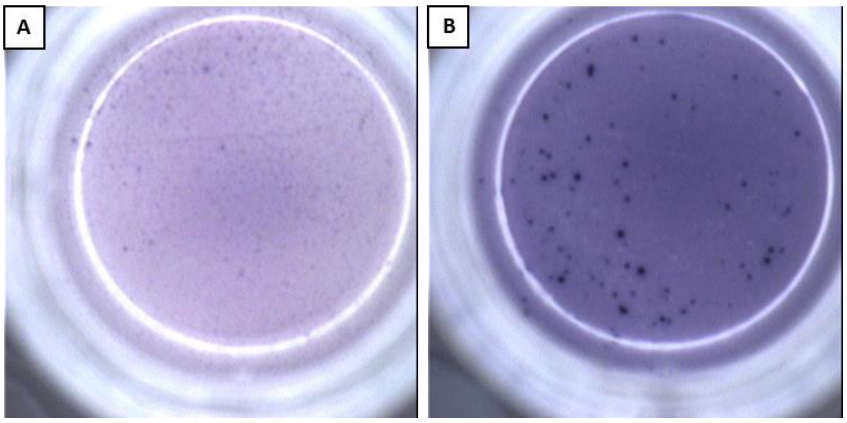

Figure 5.5. Comparison of the feline IFN-γ ELISpot assay before and after optimisation. The PBMC in well A were incubated in the assay for 18 hours and were rested prior to use in the assay, whereas the PBMC in well B were incubated in the assay for 24 hours and were not rested prior to use in the assay. The PBMC in both wells were stimulated with 10 µg/ml PHA and were plated at the same density, but they were taken from different cats.

With these steps in place, experiments began to test the ability of the peptide pools to stimulate PBMC, using cells stimulated with PHA as a positive control and unstimulated cells as a negative control. However, it was found that around 50% of the feline PBMC populations did not respond adequately to PHA (i.e. there were too few spots to distinguish the positive control from background; Figure 5.6). An experiment was therefore carried out to compare PHA (10 µg/ml) (Passlack *et al.*, 2017), concanavalin A (5 µg/ml) (Abbott *et al.*, 2012) and phorbol 12-myristate 13-acetate (PMA)/ionomycin (100/600 ng/ml) (Nordone *et al.*, 2005) as positive control mitogens (Figure 5.7). Concanavalin A was chosen as the best mitogen because PBMC responded to it more strongly than PHA, but not so strongly that spots became confluent, as was the case with PMA/ionomycin.

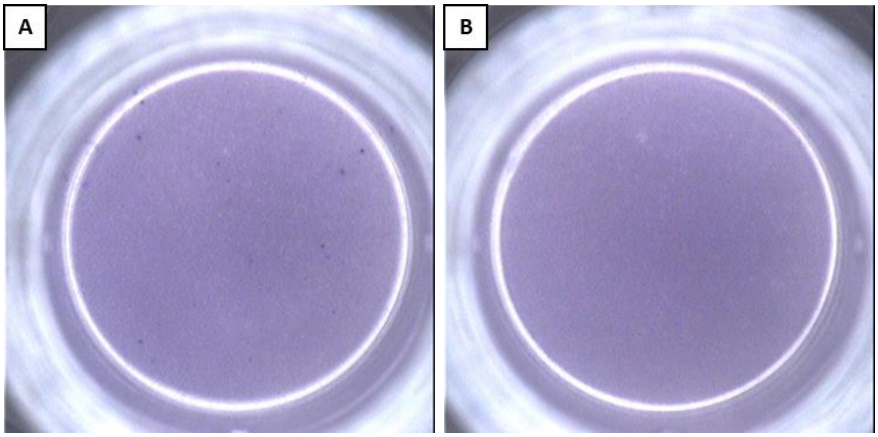

Figure 5.6. Example of a PBMC population that failed to respond to PHA in a feline IFN-γ ELISpot assay. Well A is the positive control (i.e. PBMC were treated with 10 µg/ml PHA) and well B is the negative control. The PBMC in both wells were from the same cat and were plated at the same density.

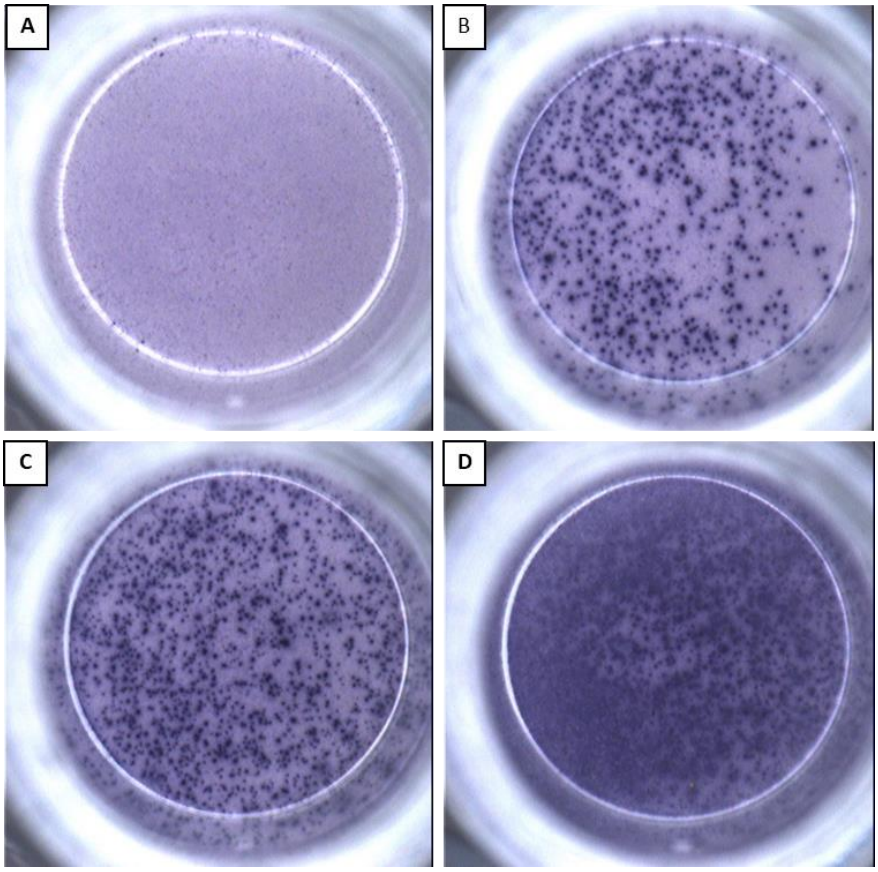

Figure 5.7. Results of a feline IFN-γ ELISpot assay carried out to determine the optimal positive control mitogen. Well A is the negative control, well B has been treated for 24 hours with PHA (10 µg/ml), well B for 24 hours with concanavalin A (5 µg/ml) and well C for 24 hours with PMA/ionomycin (100/600 ng/ml). The PBMC in all wells were from the same cat and were plated at the same density.

Another issue encountered was the high incidence of artefactual spots found under all conditions, including the negative control, that likely represented cell debris or reagent aggregation (Nordone *et al.*, 2005). The following steps were therefore taken to overcome this problem:

- Diluted reagents were passed through at 0.2 µm syringe filter before being used in the assay to remove any aggregates that may have been causing artefactual spots.
- A washing step using deionised H_2O was incorporated after cell incubation in order to lyse the cells and remove them from the plate, overcoming the problem of cell debris.

The incidence of artefactual spots was reduced to an acceptable level, but it was decided that spots would be counted by eye due to difficulty finding an ELISpot reader setting that reliably discriminated genuine from artefactual spots, despite attempts to do so.

5.2.4 Response of PBMC to peptide pools on ELISpot assay

Having optimised the feline IFN-γ ELISpot assay, PBMC from 51 seropositive cats were analysed for spot forming cells (SFC) per 10^6 cells in response to each peptide pool (peptides 28, 30 and 33: isoleucine peptides, peptides 29, 31 and 32: threonine peptides, peptides 19-46: HR1 peptides; section 5.2.1). The PBMC from each cat were counted and divided evenly into ten, to be incubated, in duplicate, with concanavalin A (positive control), culture medium only (negative control), isoleucine peptides, threonine peptides or HR1 peptides. Samples from many cats were run through the assay simultaneously. Figure 5.8 shows the typical setup of an ELISpot assay in this project.

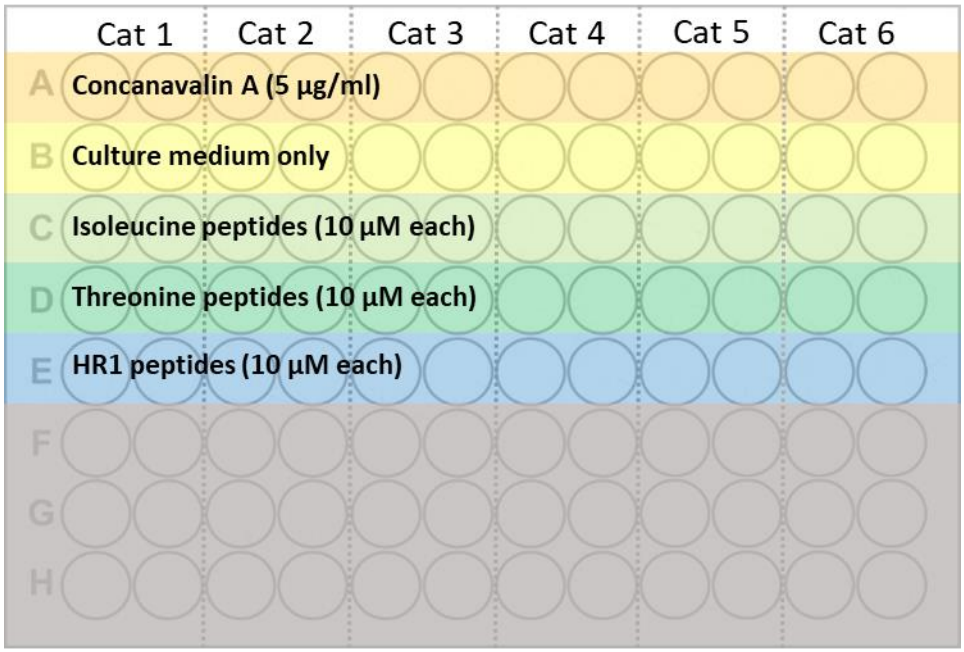

Figure 5.8. A typical 96 well plate plan for the ELISpot assay. Concanavalin A (positive control), culture medium (negative control) or a peptide pool was added to PBMC at the concentrations specified before incubation for 24 hours.

Being a larger peptide pool, HR1 peptides were included to see whether a greater response was elicited when more potential T cell epitopes were included. The data points of SFC per 10^6 cells were not normally distributed. The median value of SFC per 10^6 cells using the peptide pools for stimulation was 0 (range: 0 to 43) for the isoleucine peptides, 0 (range: 0 to 29) for the threonine peptides and 0 (range: 0 to 36) for the HR1 peptides (Figure 5.9).

This result showed that there was no median response to any peptide pool above background, likely because any response seen using the PBMC from an individual cat was obscured by the large number of PBMC populations that failed to respond to any peptide pool. The SFC per 10^6 cells for the 31 cats that responded above background to at least one peptide pool were therefore analysed separately (Figure 5.10). These data were not normally distributed. A statistically significant difference was identified between the median PBMC response to the three peptide pools (p=0.025),

with the response to the HR1 peptides being significantly higher than to both the isoleucine (p=0.001) and threonine (p<0.001) peptides.

In order to explore any difference between the responses to the isoleucine and threonine peptides, a third analysis was carried out only on the 17 cats that showed a response above background to either of these peptide pools (Figure 5.11). The data were not normally distributed. The median value of SFC per 10^6 cells using the isoleucine peptides was 3 (range: 0 to 43) and using the threonine peptides was 2 (range: 0 to 29), but this difference was not statistically significant (p=0.461). Of these 17 cats, nine (53%) showed a greater response to the isoleucine peptides, six (35%) showed a greater response to the threonine peptides and two (12%) showed an equal response to both. The 17 cats represented here had an extremely similar distribution of breeds, sex and age groups to the 51 cats originally analysed, therefore no unique factors were identified that could be associated with their response to isoleucine or threonine peptides.

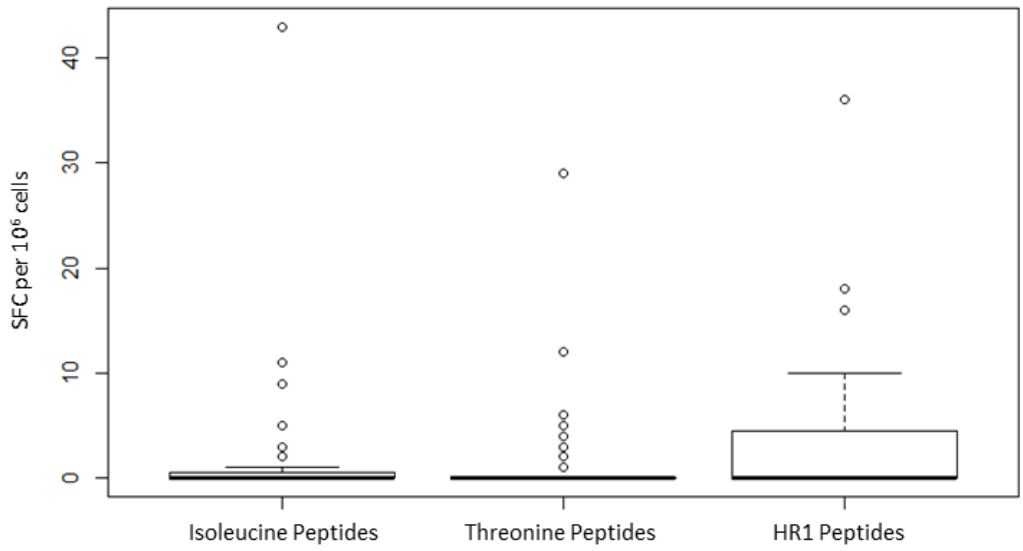

Figure 5.9. The response of feline PBMC to each peptide pool. The PBMC from 51 cats were incubated with isoleucine, threonine or HR1 peptide pools in a feline IFN-γ ELISpot assay and the number of SFC per 10^6 cells using each peptide pool were counted. The data are represented as a box plot, with the box representing the interquartile range, the thick black line representing the median, the whisker representing the maximum value excluding outliers and the circles representing outliers.

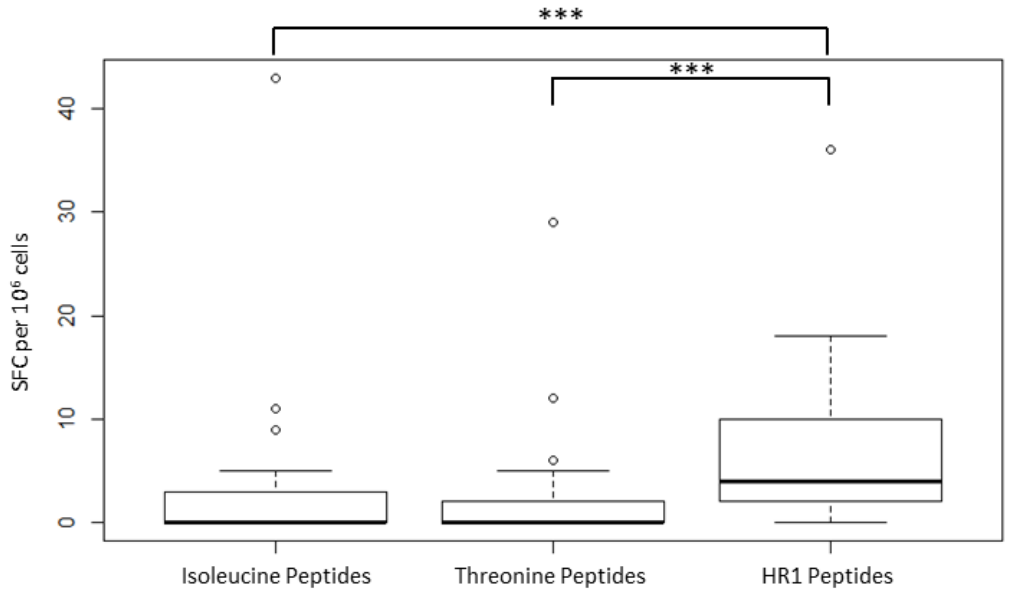

Figure 5.10. The response of feline PBMC to each peptide pool, of the cats that showed a response above background to at least one peptide pool. PBMC from 31 cats were incubated with isoleucine, threonine or HR1 peptide pools in a feline IFN-γ ELISpot assay and the number of SFC per 10^6 cells using each peptide pool were counted. The data are represented as a box plot, with the box representing the interquartile range, the thick black line representing the median, the whiskers representing the minimum and maximum values excluding outliers and the circles representing outliers. *** $p \leq 0.001$.

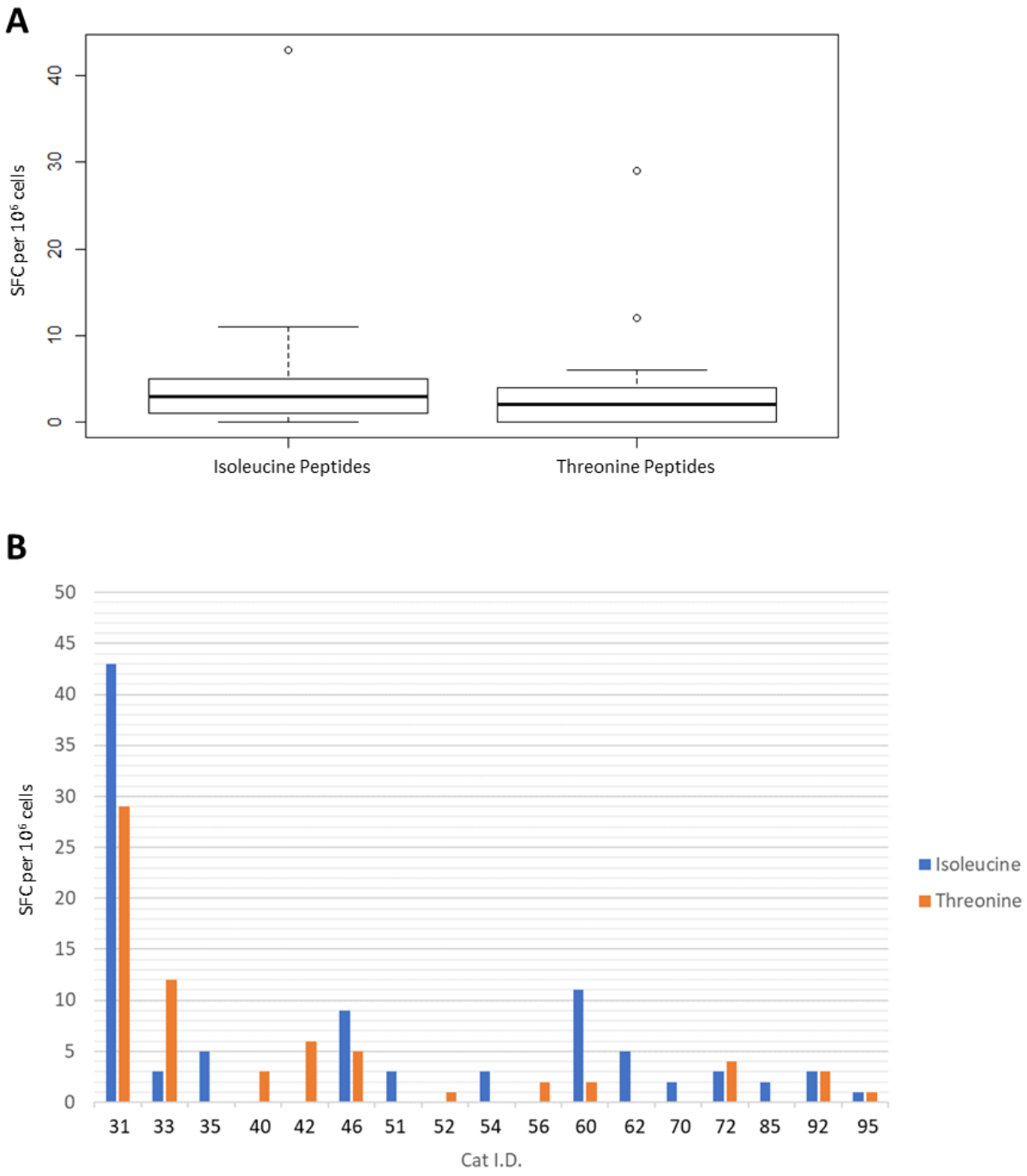

Figure 5.11. The response of feline PBMC to isoleucine or threonine peptides, of the cats that showed a response above background to either peptide pool. PBMC from 17 cats were incubated with isoleucine or threonine peptides in a feline IFN-γ ELISpot assay and the number of SFC per 10^6 cells using each peptide pool were counted;
A: The data are represented as a box plot, with the box representing the interquartile range, the thick black line representing the median, the whisker representing the minimum and maximum values excluding outliers and the circles representing outliers;
B: The response of PBMC to isoleucine or threonine peptide pools is compared for each individual cat.

5.3 Discussion

This chapter describes the optimisation and use of a feline IFN-γ ELISpot assay to assess the immunogenicity of Type 1 FCoV-derived peptides. Knowledge of the viral epitopes and variants thereof that are best suited to stimulating a selectively cellular immune response in the host would form a strong basis for rational vaccine design.

The first aim of this project was to optimise a feline IFN-γ ELISpot assay in order to test FCoV peptides for their ability to stimulate T cells. Though the use of a feline IFN-γ ELISpot has been reported (Satoh *et al.*, 2011a, Nordone *et al.*, 2005, Abbott *et al.*, 2012, Gutierrez, 2008), it is a relatively novel technique and the published literature is often in disagreement regarding the optimal methods to use. This project used a commercially available feline IFN-γ ELISpot kit containing capture and detection antibodies that are known to function together in this kind of assay at defined concentrations, but other aspects of the assay, including the length of time for which cells should be incubated, had to be optimised. Other conditions were also tested to increase spot intensity and/or reduce background colouration based on published literature and the plate manufacturer's advice.

An optimised protocol was finally deduced that produced consistent and reliable results using control mitogens for PBMC stimulation. However, one aspect of the assay that was not optimised in this project was the dose of peptide. A final concentration of 10 µM (approximately 15 µg/ml) of each peptide was chosen based on previous studies performed in the laboratory, though concentrations from 10 µg/ml (Abbott *et al.*, 2012) to 30 µg/ml (Satoh *et al.*, 2011a) have been described in the literature. As a dose of 10 µM seemed to produce adequate results (i.e. the dose produced SFC in some PBMC populations but not so many SFC as to render counting difficult), different concentrations were not trialled in this project. However, it is possible that 10 µM is not the optimal concentration, and that the frequency of SFC could be higher with a different concentration. Future work exploring varied doses of peptide would be useful.

The second aim of this project was to ascertain whether the isoleucine to threonine substitution at amino acid position 1108, located in the HR1 region of the FCoV S2 protein, alters the T cell response to this epitope, by testing variant peptides with PBMC in a feline IFN-γ ELISpot assay. Using the optimised protocol, PBMC from 51 cats were successfully tested with the isoleucine and threonine peptides, as well as a larger peptide pool covering the entire HR1 region (HR1 peptides). The response to the HR1 peptides was found to be significantly higher than the response to the isoleucine (p=0.001) and threonine (p<0.001) peptides when PBMC from the 31 cats who responded above background to any of the peptide conditions were analysed separately. By excluding non-

responders from the analysis in this way, the problem of a large number of zero data lowering the median could be overcome.

A greater response to HR1 than isoleucine or threonine peptides was expected as the larger HR1 pool contained more potential T cell epitopes, which increased the likelihood of a T cell response. Since each peptide was used at 10 µM and there were many more peptides in the HR1 than other pools, the overall concentration of peptide in the HR1-treated wells would have been higher. This should not affect the result since T cells should only become activated if they recognise specific epitopes, but it is possible that peptides in general provided a non-specific stimulatory effect on T cells, and so a higher concentration of peptides led to more non-specific stimulation. Inclusion of a peptide control to which cats are known to not respond, at the same concentration as the HR1 peptides, would rule out this effect.

Though a good indicator that at least some of the peptides were immunogenic, this analysis gave no insight into whether the I1108T mutation affects the T cell response to this particular epitope. Further analysis was therefore carried out only on the 17 cats who responded above background to either the threonine or isoleucine peptides; this showed a higher median response to the isoleucine compared to the threonine peptides, but this was not statistically significant. To overcome the possible issue of opposite trends in individual responses to isoleucine or threonine peptides offsetting each other, the response of each individual cat to the two peptide pools was compared. It was found that 53% of cats showed a greater response to the isoleucine than threonine peptides (compared to 35% for the reverse). Taken together, these results suggest that the threonine epitope variant may be less effective at stimulating T cells, which would support the hypothesis that FCoV with the I1108T substitution is less able to elicit an adequate cellular immune response in the cat. A larger sample size is required to further investigate this hypothesis and trend.

A feature of the ELISpot assays run in this project was that the frequency of SFC was generally very low. Actual counts ranged from zero to ten, though by calculating the number of SFC per 10^6 cells these counts could be amplified and compared side by side despite differing numbers of cells in the assay. A simple explanation as to why spot frequency was so low is that the number of cells used was low. Nordone *et al.* (2005) recommend 500,000 cells per well for an optimal immune response when performing feline cytokine ELISpot assays, but this was rarely achieved because too few PBMC (typically ~3,000,000) were extracted from the small surplus blood volumes available. This is a problem intrinsic to using clinical samples from cats; it is near impossible to have residual samples of 1.5 ml or more of blood in a clinical setting. Another way to improve PBMC yield from the small amount of blood available could be to trial different anticoagulants with the blood samples; EDTA

was used in this project because it is commonly used in veterinary practice therefore widely available, but Nordone *et al.* (2005) suggest that EDTA could impair cytokine induction, and instead suggest using acid citrate dextrose or sodium citrate.

Another explanation for the low frequency of spots is that some of the peptides were not immunogenic. Though the isoleucine and threonine peptides covered the amino acid sequence of the major T cell epitope described by Satoh *et al.* (2011a), this project used 15-mer peptides compared to the 20-mers used in Satoh's experiments. The reason for using 15-mers in the current project is that they should be capable of recognition by both $CD4^+$ and $CD8^+$ T cells, whereas 20-mers would likely only be recognised by $CD4^+$ T cells (Abbott *et al.*, 2012). Thus it was hoped that the 15-mers would have a better chance of eliciting a T cell response. However, it cannot be assumed that the effect found by Satoh *et al.* (2011a) would also be observed with 15-mer peptides. In future investigations, the peptides could also be designed as 20-mers to allow comparison studies.

Peptide pools, potentially containing many T cell epitopes, are commonly used in ELISpot assays as a screening tool to broadly identify regions of a protein that stimulate a T cell response. If a response is identified to a peptide pool, individual peptides can then be used in the assay to further localise the epitope(s). When a single peptide is identified as containing a T cell epitope, the specific amino acid sequence functioning as an epitope can be identified by dividing that peptide into smaller pieces or using different sequence overlaps of the peptide (Abbott *et al.*, 2012). This project, though concerned with a specific T cell epitope, used overlapping peptide pools because it is unknown exactly what sequence the epitope spans and how it is best presented to T cells. The HR1 peptide pool was used, in effect, as a positive control in this project; being a larger peptide pool and thus potentially containing more T cell epitopes, HR1 should have a higher likelihood of eliciting a T cell response. Additionally, Satoh *et al.* (2011a) identified HR1 as a major source of T cell epitopes, with four of 14 peptides covering the region stimulating a significant IFN-γ response in the study. However, HR1 peptides still only cover a small region of an FCoV protein, and a lack of response to them does not mean that the cat would be unable to mount a T cell response against another part of the virus. A stimulus such as whole inactivated virus (Satoh *et al.*, 2011a) may have been a better control for the individual's T cell response to FCoV than the HR1 peptides used in the current project.

The number of SFC per 10^6 cells was calculated from the raw counts in order to compare results between different cats, but this is potentially problematic. Since far fewer than 10^6 cells were seeded per well, a difference of one spot between wells was amplified to look like a very large difference, which may have led to placing more significance on a result than was warranted. This is a

particular hazard when spots are counted by eye, as, despite extensive efforts in the current project by the operator, such counts are always somewhat subjective. This problem could be overcome by only using samples containing PBMC above a threshold number in order to prevent over-amplification of results, and by using the ELISpot reader to count spots automatically. Another solution could be to calculate the number of SFC per a lower number of cells. This would prevent over-amplification but could lead to a situation where many of the resulting data are very low and difficult to distinguish from each other.

The PBMC from some cats were not used in the ELISpot assay, largely because of issues around the coordination of clinical sample submission. It was decided to recruit samples by advertising to breeders due to the low number of samples that were obtained through the original recruitment strategy. However, involvement of the breeders greatly increased the number of samples submitted, introducing difficulties in coordinating sample arrival in the laboratory and ensuring that they could all be used in the ELISpot assay. Freezing of extracted PBMC was initially trialled but it was found that this markedly decreased the viable cell count. It was also specified to breeders and veterinarians that submitted blood samples must be at least 1 ml in volume and delivered within two days of collection from the cat. However, some samples did not meet these requirements and it was therefore difficult or impossible to extract enough viable PBMC. In total, 12 samples were unsuitable for use in the assay; three were too small a volume, one arrived after the project had finished and eight were unsuitable for reasons unrelated to volume or time (four were haemolysed and four were clotted). Twenty samples had a smaller volume than was specified but sufficient PBMC could be extracted to permit their use in the assay.

Seroprevalence amongst the 'breeder' cohort of cats was 92%, compared to just 7% for the 'non-breeder' cohort. By recruiting samples from breeders, not only were almost all PBMC used in the ELISpot assay from young, unneutered pedigrees, but the breeders who took part were also not likely to be randomly selected in terms of FCoV infection status. Those who chose to participate in the study by obtaining free of charge FCoV serology testing for their cats were likely to have an interest in FCoV, possibly because of problems with FIP or previous positive results on FCoV serology. Though the population in this project was not representative of the wider cat population, recruiting in this way meant that most cats were FCoV seropositive so their PBMC could be used to measure T cell responses to FCoV peptides. The original recruitment strategy may have led to a more representative population but would have also meant that most samples could not have likely been used in the ELISpot assay if more had been derived from seronegative cats.

An assumption made in this project is that seropositive cats had been exposed to Type 1 FCoV. This is a reasonable assumption since the vast majority of FCoV field strains in the UK are Type 1 (Addie *et al.*, 2003) but there is a chance that some of the cats could have been infected with Type 2 FCoV. The indirect immunofluorescent antibody test used in this project to measure FCoV antibody titre recognises antibodies to both Type 1 and Type 2 FCoV so does not distinguish between the serotypes. This is an important factor, since this project looked at T cell epitopes situated in the FCoV S2 protein; a region that would vary considerably between Type 1 and Type 2 FCoV. However, methods to distinguish the two serotypes (e.g. RT-PCR and sequencing) rely on direct detection of virus, and seropositivity does not always equate to active virus shedding (Addie and Jarrett, 2001), so it may have been problematic to identify any infecting FCoV serotype in the cats. Although such a step could be introduced in the future, it is worth remembering that Type 2 FCoV is rare in the UK (Addie *et al.*, 2003).

In order to learn more about the impact of the FCoV I1108T substitution on the T cell response, it would be useful to run more PBMC samples from a wider range of cat breeds, age groups and sexes through the optimised IFN-γ ELISpot assay, increasing sample size and therefore the power to detect a significant difference in the response to the isoleucine and threonine peptides, if one exists. It would also be useful to gather more information about the cats such as existing morbidities that could affect T cell function, for example feline immunodeficiency virus.

Though no significant difference was found in the ability of the isoleucine and threonine peptide variants to stimulate T cells, an optimised ELISpot assay has been developed in this study that could be used in the future to measure the immunogenicity of T cell epitopes and their variants, and therefore guide rational vaccine design. Thinking into the future, a feline IFN-γ ELISpot assay could be a useful tool for assessing cellular immunity to FCoV post-vaccination (Abbott *et al.*, 2012, Slota *et al.*, 2011). Knowing the importance of the cellular immune response in overcoming FIP challenge, there is also the potential to use an IFN-γ ELISpot as a prognostic indicator in cats with FIP or, given the recent breakthroughs in FIP treatment with GC376 (Pedersen *et al.*, 2017) and GS-441524 (Pedersen *et al.*, 2019), as a way of assessing treatment success in cats with FIP.

6 Final perspectives and future work

The overall aim of this project was to establish a reverse genetic system for Type 1 FCoV. Such a system constitutes a powerful tool for better understanding the molecular pathogenesis of FCoV and the development of FIP, as it allows targeted mutations to be introduced into the FCoV genome and the effects of these mutations on the virus phenotype to be studied. Reverse genetic systems have also been used to generate vaccine candidates, as the platform lends itself to attenuating viruses in a rational and strategic way (Stobart and Moore, 2014).

At the beginning of this project, the only Type 1 FCoV reverse genetic system was based on the 'Black' virus strain (Tekes *et al.*, 2008). 'Black' is technically a Type 1 FCoV but is so adapted to cell culture it can no longer be said to represent field strains of the virus (Pedersen, 2009). This project therefore endeavored to achieve something novel by building a reverse genetic system based on a clinically relevant field strain of Type 1 FCoV. Three years into the project, Ehmann *et al.* (2018) achieved this aim and published the first report of a Type 1 FCoV reverse genetic system based on a field strain of the virus. This groundbreaking study demonstrated that cells transfected with full-length RNA encoding this recombinant Type 1 FCoV produced viral particles that, while not infectious *in vitro*, resulted in productive infection when given to cats.

During this project, full-length constructs representing a field strain of Type 1 FCoV, chimeric Type 1 FCoV with a Type 2 FCoV S gene and an FCoV replicon were designed and built through *in vitro* ligation. These constructs were transcribed and the latter two were transfected into cells, but no detectable virus or replicon-expressing cells were made. Infectious virus was likely not recovered from cells transfected with the chimeric construct because the construct was truncated, though unfortunately this was not discovered until after the experiment was carried out. The likely reasons that replicon-expressing cells could not be established are discussed in depth in chapter 3.3, but briefly include a non-synonymous mutation that was unavoidably introduced into the ORF1a gene during amplification in bacterial culture, and issues with the transfection method; specifically, using chemical transfection rather than electroporation.

It would be worthwhile to persevere with the constructs developed in this project, because they are based on a different strain of FCoV from that used by Ehmann *et al.* (2018) so would contribute to a fuller understanding of the virus. Additionally, an FCoV replicon has not been reported before in the literature so, if established, would be a completely novel tool for exploring the impact of virus replication on cells and rapidly screening antivirals targeting the non-structural parts of the virus. To move forward with this part of the project, the full-length constructs could be cloned into vaccinia

virus vectors. This would overcome the problem experienced in this project of having a limited amount of RNA to work with, and the mutation in the ORF1a gene could be repaired.

Unlike Type 2 FCoV, a CER for Type 1 FCoV has not been identified and there are no available cell lines in which field strains of the virus can be propagated. This is a major hurdle in FCoV research, because Type 1 FCoV is by far the most abundant serotype in nature (Benetka *et al.*, 2004, Addie *et al.*, 2003, Li *et al.*, 2018) and therefore the most biologically relevant. Ehmann *et al.* (2018) circumvented this problem by inoculating cats with their recombinant Type 1 FCoV, but the development of a system to recover recombinant virus (as well as field strains of Type 1 FCoV) *in vitro* would represent a huge leap forward in FCoV research. Not only is it more ethical and straightforward to characterise recombinant mutant viruses in cells rather than in cats, a cell-based virus replication system would facilitate the dissection of the steps that potentially lead to FIP without the complicating factors of host and environment.

That is why an important second aim of this project was to identify a CER for Type 1 FCoV. With knowledge of a receptor that this serotype uses to gain entry to its target cells, it would be possible to engineer that receptor into a cell line, rendering it permissive to infection with Type 1 FCoV and therefore overcoming an obstacle that has impeded FCoV research for many years. Prior to the start of this project, Desmarets *et al.* (2013) developed an immortalised feline intestinal epithelial cell line that was able to support the propagation of field strains of Type 1 FCoV. However, the group were unwilling to share this cell line with the scientific community, and it was not possible to replicate their results in this study despite a number of attempts.

As well as a useful tool in itself to support virus replication, the intestinal epithelial cell line previously reported by Desmarets *et al.* (2013) would have been invaluable to this project for the identification of a Type 1 FCoV CER. The approach taken in this project to identify the CER involved incubating a chimeric 'bait protein', comprising the FCoV S1 (receptor binding) domain and human IgG Fc, with a lysate of a cell type permissive to infection with the virus (and therefore bearing a CER for the virus). The bait protein was then purified *via* its IgG Fc portion, and the protein to which it had bound characterised using mass spectrometry and proteomics. This method was first used to identify a CER for MERS-CoV (Raj *et al.*, 2013), and was used in this project to 'identify' APN as a CER for Type 2 FCoV. The latter identification was straightforward and achieved with minimal optimisation, suggesting that the methodology was sound and a Type 1 FCoV CER could have been easily identified given a suitable cell line to work with.

In the absence of the use of the aforementioned intestinal cell line, two alternative cell lines, feline PBMC and feline monocyte-derived cells, were screened for expression of a Type 1 FCoV CER with no

success. A breakthrough came when lysates of feline intestinal organoids, 3D cultures derived from stem cells that self-organise into 'mini guts', were incubated with the bait proteins, and HSPA1A was found to be significantly enriched with Type 1 bait protein. This interaction was confirmed when the Type 1 bait protein was found to precipitate HSPA1A from pHSPA1A-transfected cells. Interestingly, HSPA1A is a plausible CER for Type 1 FCoV because it has been found located at the plasma membrane in a human cell line (Bilog *et al.*, 2019) and is a putative CER for a range of other viruses including intestinal rotavirus (Perez-Vargas *et al.*, 2006). An urgent next step for validating HSPA1A as a CER for Type 1 FCoV would be to test whether resistant cells are rendered infectable by Type 1 FCoV when they are made to express HSPA1A. If this is the case, HSPA1A could confidently be identified as a CER for Type 1 FCoV and work could begin to stably express this protein in, for example, CrFK cells, creating a cell line in which field strains of the virus can be propagated.

In the absence of the definitive identification of a Type 1 FCoV CER, feline intestinal organoids could still support *in vitro* propagation of Type 1 viruses. The utility of intestinal organoids for this purpose was demonstrated by Ettayebi *et al.* (2016), when the group used these cultures to grow HuNoVs *in vitro* for the first time. This was achieved through the development of monolayers: organoid cultures grown in 2D rather than 3D, to allow access to the apical (receptor-bearing) surface of the cell (van der Hee *et al.*, 2018). Intestinal organoids are more difficult to work with than most immortalised cell lines, but they more closely recapitulate the natural site of infection of FCoV so may be more biologically relevant. Organoids also carry the genetic signature of the individual from which they were derived (Ramani *et al.*, 2018), suggesting that they may offer the potential to explore virus-host interactions *in vitro*. Development of feline intestinal organoid monolayers would be the next step in exploring this avenue of research.

The purpose of this project was to build a tool that would allow researchers to gain a better understanding of how FIP develops from FCoV, and rationally design vaccines to prevent this disease. To this end, a feline IFN-γ ELISpot assay was first optimised, then used to measure the immunogenicity of peptides representing a region of the FCoV S2 protein with and without an amino acid substitution. This substitution, identified by Lewis *et al.* (2015), was found in all FCoVs from the tissues of cats with FIP but no FCoVs from the faeces of healthy cats. Sitting within a major T cell epitope for Type 1 FCoV as identified by Satoh *et al.* (2011a), it was hypothesised that the substitution could influence the immune response of the infected cat and therefore its clinical course and outcome. This project found that peptides representing the FIP-related FCoV variant were slightly less immunogenic, which would support the hypothesis. However, the difference between the two variants was not statistically significant. Beyond investigating the question of this specific substitution, it was worthwhile to optimise and gain experience of working with the assay.

The IFN-γ ELISpot assay is an established method to assess cellular immunity (Slota *et al.*, 2011, Schmittel *et al.*, 2001), which is thought to be an important factor in preventing the development of FIP from FCoV (Pedersen, 2014b). Any successful vaccine against FIP will likely have to be capable of stimulating a cellular immune response in the cat, and a feline IFN-γ ELISpot assay provides a straightforward and relevant way of testing this capability in the future.

Although a reverse genetic system for Type 1 FCoV could not be established within the time constraints of this project, the methods that have been optimised and the constructs that have been built provide a solid foundation for a future project to achieve this aim. Similarly, a cell line in which to propagate Type 1 FCoV could not be developed within the time constraints of this project, but the promising lead on a candidate receptor is ready to be followed up in a future project. This will hopefully result in the definitive identification of a Type 1 FCoV CER and a way of propagating field strains of Type 1 FCoV *in vitro.*

ABBOTT, J. R., PU, R., COLEMAN, J. K. & YAMAMOTO, J. K. 2012. Utilization of feline ELISPOT for mapping vaccine epitopes. *Methods Mol Biol,* 792**,** 47-63.

ABCAM. 2016. *ELISPOT troubleshooting tips* [Online]. Available: http://www.abcam.com/index.html?pageconfig=resource&rid=11457 [Accessed 25 February 2016].

ACAR, D. D., OLYSLAEGERS, D. A., DEDEURWAERDER, A., ROUKAERTS, I. D., BAETENS, W., VAN BOCKSTAEL, S., DE GRYSE, G. M., DESMARETS, L. M. & NAUWYNCK, H. J. 2016. Upregulation of endothelial cell adhesion molecules characterizes veins close to granulomatous infiltrates in the renal cortex of cats with feline infectious peritonitis and is indirectly triggered by feline infectious peritonitis virus infected monocytes in vitro. *J Gen Virol*.

ADDIE, D. 2012. *Feline Coronavirus Infections,* St. Louis, MO, Elsevier/Saunders.

ADDIE, D., BELAK, S., BOUCRAUT-BARALON, C., EGBERINK, H., FRYMUS, T., GRUFFYDD-JONES, T., HARTMANN, K., HOSIE, M. J., LLORET, A., LUTZ, H., MARSILIO, F., PENNISI, M. G., RADFORD, A. D., THIRY, E., TRUYEN, U. & HORZINEK, M. C. 2009. Feline infectious peritonitis. ABCD guidelines on prevention and management. *J Feline Med Surg,* 11**,** 594-604.

ADDIE, D., HOUE, L., MAITLAND, K., PASSANTINO, G. & DECARO, N. 2019. Effect of cat litters on feline coronavirus infection of cell culture and cats. *J Feline Med Surg***,** 1098612X19848167.

ADDIE, D. D. & JARRETT, O. 1992. A study of naturally occurring feline coronavirus infections in kittens. *Vet Rec,* 130**,** 133-7.

ADDIE, D. D. & JARRETT, O. 2001. Use of a reverse-transcriptase polymerase chain reaction for monitoring the shedding of feline coronavirus by healthy cats. *Vet Rec,* 148**,** 649-53.

ADDIE, D. D., LE PODER, S., BURR, P., DECARO, N., GRAHAM, E., HOFMANN-LEHMANN, R., JARRETT, O., MCDONALD, M. & MELI, M. L. 2015. Utility of feline coronavirus antibody tests. *J Feline Med Surg,* 17**,** 152-62.

ADDIE, D. D., PALTRINIERI, S., PEDERSEN, N. C. & SECONG INTERNATIONAL FELINE CORONAVIRUS/FELINE INFECTIOUS PERITONITIS, S. 2004. Recommendations from workshops of the second international feline coronavirus/feline infectious peritonitis symposium. *J Feline Med Surg,* 6**,** 125-30.

ADDIE, D. D., SCHAAP, I. A., NICOLSON, L. & JARRETT, O. 2003. Persistence and transmission of natural type I feline coronavirus infection. *J Gen Virol,* 84**,** 2735-44.

ADDIE, D. D., TOTH, S., MURRAY, G. D. & JARRETT, O. 1995. Risk of feline infectious peritonitis in cats naturally infected with feline coronavirus. *Am J Vet Res,* 56**,** 429-34.

ALAGAILI, A. N., BRIESE, T., MISHRA, N., KAPOOR, V., SAMEROFF, S. C., BURBELO, P. D., DE WIT, E., MUNSTER, V. J., HENSLEY, L. E., ZALMOUT, I. S., KAPOOR, A., EPSTEIN, J. H., KARESH, W. B., DASZAK, P., MOHAMMED, O. B. & LIPKIN, W. I. 2014. Middle East respiratory syndrome coronavirus infection in dromedary camels in Saudi Arabia. *MBio,* 5**,** e00884-14.

ALMAZAN, F., GALAN, C. & ENJUANES, L. 2004. The nucleoprotein is required for efficient coronavirus genome replication. *J Virol,* 78**,** 12683-8.

ALMAZAN, F., SOLA, I., ZUNIGA, S., MARQUEZ-JURADO, S., MORALES, L., BECARES, M. & ENJUANES, L. 2014. Coronavirus reverse genetic systems: infectious clones and replicons. *Virus Res,* 189**,** 262-70.

ALSAADI, E. A. J. & JONES, I. M. 2019. Membrane binding proteins of coronaviruses. 14**,** 275-286.

ANIS, E. A., DHAR, M., LEGENDRE, A. M. & WILKES, R. P. 2016. Transduction of hematopoietic stem cells to stimulate RNA interference against feline infectious peritonitis. *J Feline Med Surg*.

BALINT, A., FARSANG, A., SZEREDI, L., ZADORI, Z. & BELAK, S. 2014. Recombinant feline coronaviruses as vaccine candidates confer protection in SPF but not in conventional cats. *Vet Microbiol,* 169**,** 154-62.

BALINT, A., FARSANG, A., ZADORI, Z., HORNYAK, A., DENCSO, L., ALMAZAN, F., ENJUANES, L. & BELAK, S. 2012. Molecular characterization of feline infectious peritonitis virus strain DF-2 and studies of the role of ORF3abc in viral cell tropism. *J Virol,* 86**,** 6258-67.

BANK-WOLF, B. R., STALLKAMP, I., WIESE, S., MORITZ, A., TEKES, G. & THIEL, H. J. 2014. Mutations of 3c and spike protein genes correlate with the occurrence of feline infectious peritonitis. *Vet Microbiol,* 173**,** 177-88.

BANTSCHEFF, M., SCHIRLE, M., SWEETMAN, G., RICK, J. & KUSTER, B. 2007. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem,* 389**,** 1017-31.

BARIC, R. S., NELSON, G. W., FLEMING, J. O., DEANS, R. J., KECK, J. G., CASTEEL, N. & STOHLMAN, S. A. 1988. Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. *J Virol,* 62**,** 4280-7.

BARKER, E. N., STRANIERI, A., HELPS, C. R., PORTER, E. L., DAVIDSON, A. D., DAY, M. J., KNOWLES, T., KIPAR, A. & TASKER, S. 2017. Limitations of using feline coronavirus spike protein gene mutations to diagnose feline infectious peritonitis. *Vet Res,* 48**,** 60.

BARTFELD, S. & CLEVERS, H. 2015. Organoids as Model for Infectious Diseases: Culture of Human and Murine Stomach Organoids and Microinjection of Helicobacter Pylori. *J Vis Exp*.

BEALL, A., YOUNT, B., LIN, C. M., HOU, Y., WANG, Q., SAIF, L. & BARIC, R. 2016. Characterization of a Pathogenic Full-Length cDNA Clone and Transmission Model for Porcine Epidemic Diarrhea Virus Strain PC22A. *MBio,* 7**,** e01451-15.

BELOUZARD, S., MILLET, J. K., LICITRA, B. N. & WHITTAKER, G. R. 2012. Mechanisms of coronavirus cell entry mediated by the viral spike protein. *Viruses,* 4**,** 1011-33.

BENBACER, L., KUT, E., BESNARDEAU, L., LAUDE, H. & DELMAS, B. 1997. Interspecies aminopeptidase-N chimeras reveal species-specific receptor recognition by canine coronavirus, feline infectious peritonitis virus, and transmissible gastroenteritis virus. *J Virol,* 71**,** 734-7.

BENETKA, V., KUBBER-HEISS, A., KOLODZIEJEK, J., NOWOTNY, N., HOFMANN-PARISOT, M. & MOSTL, K. 2004. Prevalence of feline coronavirus types I and II in cats with histopathologically verified feline infectious peritonitis. *Vet Microbiol,* 99**,** 31-42.

BILOG, A. D., SMULDERS, L., OLIVERIO, R., LABANIEH, C., ZAPANTA, J., STAHELIN, R. V. & NIKOLAIDIS, N. 2019. Membrane Localization of HspA1A, a Stress Inducible 70-kDa Heat-Shock Protein, Depends on Its Interaction with Intracellular Phosphatidylserine. *Biomolecules,* 9.

BOETTCHER, I. C., STEINBERG, T., MATIASEK, K., GREENE, C. E., HARTMANN, K. & FISCHER, A. 2007. Use of anti-coronavirus antibody testing of cerebrospinal fluid for diagnosis of feline infectious peritonitis involving the central nervous system in cats. *J Am Vet Med Assoc,* 230**,** 199-205.

BORSCHENSKY, C. M. & REINACHER, M. 2014. Mutations in the 3c and 7b genes of feline coronavirus in spontaneously affected FIP cats. *Res Vet Sci,* 97**,** 333-40.

BOSCH, B. J., VAN DER ZEE, R., DE HAAN, C. A. & ROTTIER, P. J. 2003. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol,* 77**,** 8801-11.

BROWN, M. A., TROYER, J. L., PECON-SLATTERY, J., ROELKE, M. E. & O'BRIEN, S. J. 2009. Genetics and pathogenesis of feline infectious peritonitis virus. *Emerg Infect Dis,* 15**,** 1445-52.

BURKARD, C., VERHEIJE, M. H., WICHT, O., VAN KASTEREN, S. I., VAN KUPPEVELD, F. J., HAAGMANS, B. L., PELKMANS, L., ROTTIER, P. J., BOSCH, B. J. & DE HAAN, C. A. 2014. Coronavirus cell entry occurs through the endo-/lysosomal pathway in a proteolysis-dependent manner. *PLoS Pathog,* 10**,** e1004502.

CAVANAGH, D. P. D. 2008. *SARS- and other coronaviruses : laboratory protocols,* New York, N.Y., Humana Press.

CHA, R. S. & THILLY, W. G. 1993. Specificity, efficiency, and fidelity of PCR. *PCR Methods Appl,* 3**,** S18-29.

CHANG, H. W., DE GROOT, R. J., EGBERINK, H. F. & ROTTIER, P. J. 2010. Feline infectious peritonitis: insights into feline coronavirus pathobiogenesis and epidemiology based on genetic analysis of the viral 3c gene. *J Gen Virol,* 91**,** 415-20.

CHANG, H. W., EGBERINK, H. F., HALPIN, R., SPIRO, D. J. & ROTTIER, P. J. 2012. Spike protein fusion peptide and feline coronavirus virulence. *Emerg Infect Dis,* 18**,** 1089-95.

CHRISTIANSON, K. K., INGERSOLL, J. D., LANDON, R. M., PFEIFFER, N. E. & GERBER, J. D. 1989. Characterization of a temperature sensitive feline infectious peritonitis coronavirus. *Arch Virol,* 109**,** 185-96.

CORAPI, W. V., OLSEN, C. W. & SCOTT, F. W. 1992. Monoclonal antibody analysis of neutralization and antibody-dependent enhancement of feline infectious peritonitis virus. *J Virol,* 66**,** 6695-705.

CORNELISSEN, E., DEWERCHIN, H. L., VAN HAMME, E. & NAUWYNCK, H. J. 2009. Absence of antibody-dependent, complement-mediated lysis of feline infectious peritonitis virus-infected cells. *Virus Res,* 144**,** 285-9.

DAS, S., LAXMINARAYANA, S. V., CHANDRA, N., RAVI, V. & DESAI, A. 2009. Heat shock protein 70 on Neuro2a cells is a putative receptor for Japanese encephalitis virus. *Virology,* 385**,** 47-57.

DATE, S. & SATO, T. 2015. Mini-gut organoids: reconstitution of the stem cell niche. *Annu Rev Cell Dev Biol,* 31**,** 269-89.

DE GROOT-MIJNES, J. D., VAN DUN, J. M., VAN DER MOST, R. G. & DE GROOT, R. J. 2005. Natural history of a recurrent feline coronavirus infection and the role of cellular immunity in survival and disease. *J Virol,* 79**,** 1036-44.

DE GROOT, R. J., BAKER, S. C., BARIC, R. S., BROWN, C. S., DROSTEN, C., ENJUANES, L., FOUCHIER, R. A., GALIANO, M., GORBALENYA, A. E., MEMISH, Z. A., PERLMAN, S., POON, L. L., SNIJDER, E. J., STEPHENS, G. M., WOO, P. C., ZAKI, A. M., ZAMBON, M. & ZIEBUHR, J. 2013. Middle East respiratory syndrome coronavirus (MERS-CoV): announcement of the Coronavirus Study Group. *J Virol,* 87**,** 7790-2.

DE HAAN, C. A., HAIJEMA, B. J., SCHELLEN, P., WICHGERS SCHREUR, P., TE LINTELO, E., VENNEMA, H. & ROTTIER, P. J. 2008. Cleavage of group 1 coronavirus spike proteins: how furin cleavage is traded off against heparan sulfate binding upon cell culture adaptation. *J Virol,* 82**,** 6078-83.

DECLERCQ, J., DE BOSSCHERE, H., SCHWARZKOPF, I. & DECLERCQ, L. 2008. Papular cutaneous lesions in a cat associated with feline infectious peritonitis. *Vet Dermatol,* 19**,** 255-8.

DEDEURWAERDER, A., DESMARETS, L. M., OLYSLAEGERS, D. A., VERMEULEN, B. L., DEWERCHIN, H. L. & NAUWYNCK, H. J. 2013. The role of accessory proteins in the replication of feline infectious peritonitis virus in peripheral blood monocytes. *Vet Microbiol,* 162**,** 447-55.

DEDEURWAERDER, A., OLYSLAEGERS, D. A., DESMARETS, L. M., ROUKAERTS, I. D., THEUNS, S. & NAUWYNCK, H. J. 2014. ORF7-encoded accessory protein 7a of feline infectious peritonitis virus as a counteragent against IFN-alpha-induced antiviral response. *J Gen Virol,* 95**,** 393-402.

DELMAS, B., GELFI, J., L'HARIDON, R., VOGEL, L. K., SJOSTROM, H., NOREN, O. & LAUDE, H. 1992. Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. *Nature,* 357**,** 417-20.

DENISON, M. R., GRAHAM, R. L., DONALDSON, E. F., ECKERLE, L. D. & BARIC, R. S. 2011. Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol,* 8**,** 270-9.

DENU, J. M. 2005. Vitamin B3 and sirtuin function. *Trends Biochem Sci,* 30**,** 479-83.

DESMARETS, L. M., THEUNS, S., OLYSLAEGERS, D. A., DEDEURWAERDER, A., VERMEULEN, B. L., ROUKAERTS, I. D. & NAUWYNCK, H. J. 2013. Establishment of feline intestinal epithelial cell cultures for the propagation and study of feline enteric coronaviruses. *Vet Res,* 44**,** 71.

DESMARETS, L. M., VERMEULEN, B. L., THEUNS, S., CONCEICAO-NETO, N., ZELLER, M., ROUKAERTS, I. D., ACAR, D. D., OLYSLAEGERS, D. A., VAN RANST, M., MATTHIJNSSENS, J. & NAUWYNCK, H. J. 2016. Experimental feline enteric coronavirus infection reveals an aberrant infection

pattern and shedding of mutants with impaired infectivity in enterocyte cultures. *Sci Rep,* 6**,** 20022.
DEWERCHIN, H. L., CORNELISSEN, E. & NAUWYNCK, H. J. 2005. Replication of feline coronaviruses in peripheral blood monocytes. *Arch Virol,* 150**,** 2483-500.
DOENGES, S. J., WEBER, K., DORSCH, R., FUX, R., FISCHER, A., MATIASEK, L. A., MATIASEK, K. & HARTMANN, K. 2016a. Detection of feline coronavirus in cerebrospinal fluid for diagnosis of feline infectious peritonitis in cats with and without neurological signs. *J Feline Med Surg,* 18**,** 104-9.
DOENGES, S. J., WEBER, K., DORSCH, R., FUX, R. & HARTMANN, K. 2016b. Comparison of real-time reverse transcriptase polymerase chain reaction of peripheral blood mononuclear cells, serum and cell-free body cavity effusion for the diagnosis of feline infectious peritonitis. *J Feline Med Surg*.
DOKI, T., TAKANO, T., KAWAGOE, K., KITO, A. & HOHDATSU, T. 2016. Therapeutic effect of anti-feline TNF-alpha monoclonal antibody for feline infectious peritonitis. *Res Vet Sci,* 104**,** 17-23.
DOKI, T., TAKANO, T., KOYAMA, Y. & HOHDATSU, T. 2015. Identification of the peptide derived from S1 domain that inhibits type I and type II feline infectious peritonitis virus infection. *Virus Res,* 204**,** 13-20.
DRECHSLER, Y., ALCARAZ, A., BOSSONG, F. J., COLLISSON, E. W. & DINIZ, P. P. 2011. Feline coronavirus in multicat environments. *Vet Clin North Am Small Anim Pract,* 41**,** 1133-69.
DUARTE, A., VEIGA, I. & TAVARES, L. 2009. Genetic diversity and phylogenetic analysis of Feline Coronavirus sequences from Portugal. *Vet Microbiol,* 138**,** 163-8.
DUFOUR, D., MATEOS-GOMEZ, P. A., ENJUANES, L., GALLEGO, J. & SOLA, I. 2011. Structure and functional relevance of a transcription-regulating sequence involved in coronavirus discontinuous RNA synthesis. *J Virol,* 85**,** 4963-73.
DUNBAR, D., KWOK, W., GRAHAM, E., ARMITAGE, A., IRVINE, R., JOHNSTON, P., MCDONALD, M., MONTGOMERY, D., NICOLSON, L., ROBERTSON, E., WEIR, W. & ADDIE, D. D. 2018. Diagnosis of non-effusive feline infectious peritonitis by reverse transcriptase quantitative PCR from mesenteric lymph node fine-needle aspirates. *J Feline Med Surg***,** 1098612X18809165.
DUTHIE, S., ECKERSALL, P. D., ADDIE, D. D., LAWRENCE, C. E. & JARRETT, O. 1997. Value of alpha 1-acid glycoprotein in the diagnosis of feline infectious peritonitis. *Vet Rec,* 141**,** 299-303.
DVEKSLER, G. S., PENSIERO, M. N., CARDELLICHIO, C. B., WILLIAMS, R. K., JIANG, G. S., HOLMES, K. V. & DIEFFENBACH, C. W. 1991. Cloning of the mouse hepatitis virus (MHV) receptor: expression in human and hamster cell lines confers susceptibility to MHV. *J Virol,* 65**,** 6881-91.
DYE, C. 2006. *Molecular characterisation of type I and type II feline coronavirus genomes.* PhD, University of Bristol.
DYE, C., TEMPERTON, N. & SIDDELL, S. G. 2007. Type I feline coronavirus spike glycoprotein fails to recognize aminopeptidase N as a functional receptor on feline cell lines. *J Gen Virol,* 88**,** 1753-60.
EHMANN, R., KRISTEN-BURMANN, C., BANK-WOLF, B., KONIG, M., HERDEN, C., HAIN, T., THIEL, H. J., ZIEBUHR, J. & TEKES, G. 2018. Reverse Genetics for Type I Feline Coronavirus Field Isolate To Study the Molecular Pathogenesis of Feline Infectious Peritonitis. *MBio,* 9.
ETTAYEBI, K., CRAWFORD, S. E., MURAKAMI, K., BROUGHMAN, J. R., KARANDIKAR, U., TENGE, V. R., NEILL, F. H., BLUTT, S. E., ZENG, X. L., QU, L., KOU, B., OPEKUN, A. R., BURRIN, D., GRAHAM, D. Y., RAMANI, S., ATMAR, R. L. & ESTES, M. K. 2016. Replication of human noroviruses in stem cell-derived human enteroids. *Science,* 353**,** 1387-1393.
FEHR, A. R. & PERLMAN, S. 2015. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol,* 1282**,** 1-23.
FELTEN, S., LEUTENEGGER, C. M., BALZER, H. J., PANTCHEV, N., MATIASEK, K., WESS, G., EGBERINK, H. & HARTMANN, K. 2017a. Sensitivity and specificity of a real-time reverse transcriptase

polymerase chain reaction detecting feline coronavirus mutations in effusion and serum/plasma of cats to diagnose feline infectious peritonitis. *BMC Vet Res,* 13**,** 228.

FELTEN, S., MATIASEK, K., GRUENDL, S., SANGL, L. & HARTMANN, K. 2017b. Utility of an immunocytochemical assay using aqueous humor in the diagnosis of feline infectious peritonitis. *Vet Ophthalmol*.

FELTEN, S., MATIASEK, K., GRUENDL, S., SANGL, L., WESS, G. & HARTMANN, K. 2016. Investigation into the utility of an immunocytochemical assay in body cavity effusions for diagnosis of feline infectious peritonitis. *J Feline Med Surg*.

FELTEN, S., WEIDER, K., DOENGES, S., GRUENDL, S., MATIASEK, K., HERMANNS, W., MUELLER, E., MATIASEK, L., FISCHER, A., WEBER, K., HIRSCHBERGER, J., WESS, G. & HARTMANN, K. 2015. Detection of feline coronavirus spike gene mutations as a tool to diagnose feline infectious peritonitis. *J Feline Med Surg*.

FIELDS, B. N., KNIPE, D. M. & HOWLEY, P. M. 2013. *Fields virology,* Philadelphia, Pa., Wolters Kluwer/Lippincott Williams & Wilkins.

FISCHER, Y., RITZ, S., WEBER, K., SAUTER-LOUIS, C. & HARTMANN, K. 2011. Randomized, placebo controlled study of the effect of propentofylline on survival time and quality of life of cats with feline infectious peritonitis. *J Vet Intern Med,* 25**,** 1270-6.

FOLEY, J. E., POLAND, A., CARLSON, J. & PEDERSEN, N. C. 1997. Risk factors for feline infectious peritonitis among cats in multiple-cat environments with endemic feline enteric coronavirus. *J Am Vet Med Assoc,* 210**,** 1313-8.

GIORI, L., GIORDANO, A., GIUDICE, C., GRIECO, V. & PALTRINIERI, S. 2011. Performances of different diagnostic tests for feline infectious peritonitis in challenging clinical cases. *J Small Anim Pract,* 52**,** 152-7.

GLANSBEEK, H. L., HAAGMANS, B. L., TE LINTELO, E. G., EGBERINK, H. F., DUQUESNE, V., AUBERT, A., HORZINEK, M. C. & ROTTIER, P. J. 2002. Adverse effects of feline IL-12 during DNA vaccination against feline infectious peritonitis virus. *J Gen Virol,* 83**,** 1-10.

GOEBEL, S. J., HSUE, B., DOMBROWSKI, T. F. & MASTERS, P. S. 2004. Characterization of the RNA components of a putative molecular switch in the 3' untranslated region of the murine coronavirus genome. *J Virol,* 78**,** 669-82.

GOLOVKO, L., LYONS, L. A., LIU, H., SORENSEN, A., WEHNERT, S. & PEDERSEN, N. C. 2013. Genetic susceptibility to feline infectious peritonitis in Birman cats. *Virus Res,* 175**,** 58-63.

GONON, V., DUQUESNE, V., KLONJKOWSKI, B., MONTEIL, M., AUBERT, A. & ELOIT, M. 1999. Clearance of infection in cats naturally infected with feline coronaviruses is associated with an anti-S glycoprotein antibody response. *J Gen Virol,* 80 (Pt 9)**,** 2315-7.

GRIFFITHS, A. J. F. 2000. *An introduction to genetic analysis,* New York, W.H. Freeman.

GUNN-MOORE, D. A., GRUFFYDD-JONES, T. J. & HARBOUR, D. A. 1998. Detection of feline coronaviruses by culture and reverse transcriptase-polymerase chain reaction of blood samples from healthy cats and cats with clinical feline infectious peritonitis. *Vet Microbiol,* 62**,** 193-205.

GUTIERREZ, A. M. 2008. *Priming Immunization with a vif-deleted Feline Immunodeficiency Virus Proviral DNA Vaccine Boosted with a Killed Whole Virus Vaccine.* Doctor of Philosophy, University of California Davis.

HAIJEMA, B. J., VOLDERS, H. & ROTTIER, P. J. M. 2003. Switching Species Tropism: an Effective Way To Manipulate the Feline Coronavirus Genome. *Journal of Virology,* 77**,** 4528-4538.

HAIJEMA, B. J., VOLDERS, H. & ROTTIER, P. J. M. 2004. Live, Attenuated Coronavirus Vaccines through the Directed Deletion of Group-Specific Genes Provide Protection against Feline Infectious Peritonitis. *Journal of Virology,* 78**,** 3863-3871.

HANNEMANN, H., SUNG, P. Y., CHIU, H. C., YOUSUF, A., BIRD, J., LIM, S. P. & DAVIDSON, A. D. 2013. Serotype-specific differences in dengue virus non-structural protein 5 nuclear localization. *J Biol Chem,* 288**,** 22621-35.

HARTMANN, K. 2005. Feline infectious peritonitis. *Vet Clin North Am Small Anim Pract,* 35**,** 39-79, vi.

HARTMANN, K., BINDER, C., HIRSCHBERGER, J., COLE, D., REINACHER, M., SCHROO, S., FROST, J., EGBERINK, H., LUTZ, H. & HERMANNS, W. 2003. Comparison of different tests to diagnose feline infectious peritonitis. *J Vet Intern Med,* 17**,** 781-90.

HARUN, M. S., KUAN, C. O., SELVARAJAH, G. T., WEI, T. S., ARSHAD, S. S., HAIR BEJO, M. & OMAR, A. R. 2013. Transcriptional profiling of feline infectious peritonitis virus infection in CRFK cells and in PBMCs from FIP diagnosed cats. *Virol J,* 10**,** 329.

HARVEY, C. J., LOPEZ, J. W. & HENDRICK, M. J. 1996. An uncommon intestinal manifestation of feline infectious peritonitis: 26 cases (1986-1993). *J Am Vet Med Assoc,* 209**,** 1117-20.

HAZUCHOVA, K., HELD, S. & NEIGER, R. 2016. Usefulness of acute phase proteins in differentiating between feline infectious peritonitis and other diseases in cats with body cavity effusions. *J Feline Med Surg*.

HERREWEGH, A. A., VENNEMA, H., HORZINEK, M. C., ROTTIER, P. J. & DE GROOT, R. J. 1995. The molecular genetics of feline coronaviruses: comparative sequence analysis of the ORF7a/7b transcription unit of different biotypes. *Virology,* 212**,** 622-31.

HOHDATSU, T., IZUMIYA, Y., YOKOYAMA, Y., KIDA, K. & KOYAMA, H. 1998. Differences in virus receptor for type I and type II feline infectious peritonitis virus. *Arch Virol,* 143**,** 839-50.

HOHDATSU, T., OKADA, S., ISHIZUKA, Y., YAMADA, H. & KOYAMA, H. 1992. The prevalence of types I and II feline coronavirus infections in cats. *J Vet Med Sci,* 54**,** 557-62.

HORA, A. S., TONIETTI, P. O., TANIWAKI, S. A., ASANO, K. M., MAIORKA, P., RICHTZENHAIN, L. J. & BRANDAO, P. E. 2016. Feline Coronavirus 3c Protein: A Candidate for a Virulence Marker? *Biomed Res Int,* 2016**,** 8560691.

HSIEH, L. E. & CHUEH, L. L. 2014. Identification and genotyping of feline infectious peritonitis-associated single nucleotide polymorphisms in the feline interferon-gamma gene. *Vet Res,* 45**,** 57.

HU, C. J., CHANG, W. S., FANG, Z. S., CHEN, Y. T., WANG, W. L., TSAI, H. H., CHUEH, L. L., TAKANO, T., HOHDATSU, T. & CHEN, H. W. 2017. Nanoparticulate vacuolar ATPase blocker exhibits potent host-targeted antiviral activity against feline coronavirus. *Sci Rep,* 7**,** 13043.

HUAN, C. C., WANG, Y., NI, B., WANG, R., HUANG, L., REN, X. F., TONG, G. Z., DING, C., FAN, H. J. & MAO, X. 2015. Porcine epidemic diarrhea virus uses cell-surface heparan sulfate as an attachment factor. *Arch Virol,* 160**,** 1621-8.

HURST, K. R., YE, R., GOEBEL, S. J., JAYARAMAN, P. & MASTERS, P. S. 2010. An interaction between the nucleocapsid protein and a component of the replicase-transcriptase complex is crucial for the infectivity of coronavirus genomic RNA. *J Virol,* 84**,** 10276-88.

ICC. 2015. *Life Stages* [Online]. Available: http://icatcare.org/advice/life-stages [Accessed 23 February 2016].

ISHIDA, T., SHIBANAI, A., TANAKA, S., UCHIDA, K. & MOCHIZUKI, M. 2004. Use of recombinant feline interferon and glucocorticoid in the treatment of feline infectious peritonitis. *J Feline Med Surg,* 6**,** 107-9.

IVES, E. J., VANHAESEBROUCK, A. E. & CIAN, F. 2013. Immunocytochemical demonstration of feline infectious peritonitis virus within cerebrospinal fluid macrophages. *J Feline Med Surg,* 15**,** 1149-53.

JAIMES, J. A. & WHITTAKER, G. R. 2018. Feline coronavirus: Insights into viral pathogenesis based on the spike protein structure and function. *Virology*.

JARRETT, O., LAIRD, H. M. & HAY, D. 1973. Determinants of the host range of feline leukaemia viruses. *J Gen Virol,* 20**,** 169-75.

JOLLY, C. L. & SATTENTAU, Q. J. 2013. Attachment factors. *Adv Exp Med Biol,* 790**,** 1-23.

KASS, P., DENT, T. 1995. The Epidemiology of Feline Infectious Peritonitis in Catteries. *Feline Practice,* 23**,** 27-32.

KEDWARD-DIXON, H., BARKER, E. N., TASKER, S., KIPAR, A. & HELPS, C. R. 2019. Evaluation of polymorphisms in inflammatory mediator and cellular adhesion genes as risk factors for feline infectious peritonitis. *J Feline Med Surg***,** 1098612X19865637.

KENNEDY, M., BOEDEKER, N., GIBBS, P. & KANIA, S. 2001. Deletions in the 7a ORF of feline coronavirus associated with an epidemic of feline infectious peritonitis. *Vet Microbiol,* 81**,** 227-34.

KHACHATOORIAN, R., COHN, W., BUZZANCO, A., RIAHI, R., ARUMUGASWAMI, V., DASGUPTA, A., WHITELEGGE, J. P. & FRENCH, S. W. 2018. HSP70 Copurifies with Zika Virus Particles. *Virology,* 522**,** 228-233.

KIPAR, A., BAPTISTE, K., BARTH, A. & REINACHER, M. 2006. Natural FCoV infection: cats with FIP exhibit significantly higher viral loads than healthy infected cats. *J Feline Med Surg,* 8**,** 69-72.

KIPAR, A., BELLMANN, S., GUNN-MOORE, D. A., LEUKERT, W., KOHLER, K., MENGER, S. & REINACHER, M. 1999a. Histopathological alterations of lymphatic tissues in cats without feline infectious peritonitis after long-term exposure to FIP virus. *Vet Microbiol,* 69**,** 131-7.

KIPAR, A., BELLMANN, S., KREMENDAHL, J., KOHLER, K. & REINACHER, M. 1998a. Cellular composition, coronavirus antigen expression and production of specific antibodies in lesions in feline infectious peritonitis. *Vet Immunol Immunopathol,* 65**,** 243-57.

KIPAR, A., KOEHLER, K., BELLMANN, S. & REINACHER, M. 1999b. Feline infectious peritonitis presenting as a tumour in the abdominal cavity. *Vet Rec,* 144**,** 118-22.

KIPAR, A., KOHLER, K., LEUKERT, W. & REINACHER, M. 2001. A comparison of lymphatic tissues from cats with spontaneous feline infectious peritonitis (FIP), cats with FIP virus infection but no FIP, and cats with no infection. *J Comp Pathol,* 125**,** 182-91.

KIPAR, A., KREMENDAHL, J., ADDIE, D. D., LEUKERT, W., GRANT, C. K. & REINACHER, M. 1998b. Fatal enteritis associated with coronavirus infection in cats. *J Comp Pathol,* 119**,** 1-14.

KIPAR, A., MAY, H., MENGER, S., WEBER, M., LEUKERT, W. & REINACHER, M. 2005. Morphologic features and development of granulomatous vasculitis in feline infectious peritonitis. *Vet Pathol,* 42**,** 321-30.

KIPAR, A. & MELI, M. L. 2014. Feline infectious peritonitis: still an enigma? *Vet Pathol,* 51**,** 505-26.

KIPAR, A., MELI, M. L., BAPTISTE, K. E., BOWKER, L. J. & LUTZ, H. 2010. Sites of feline coronavirus persistence in healthy cats. *J Gen Virol,* 91**,** 1698-707.

KISS, I., POLAND, A. M. & PEDERSEN, N. C. 2004. Disease outcome and cytokine responses in cats immunized with an avirulent feline infectious peritonitis virus (FIPV)-UCD1 and challenge-exposed with virulent FIPV-UCD8. *J Feline Med Surg,* 6**,** 89-97.

KLEPFER, S., REED, A. P., MARTINEZ, M., BHOGAL, B., JONES, E. & MILLER, T. J. 1995. Cloning and expression of FECV spike gene in vaccinia virus. Immunization with FECV S causes early death after FIPV challenge. *Adv Exp Med Biol,* 380**,** 235-41.

KUMMROW, M., MELI, M. L., HAESSIG, M., GOENCZI, E., POLAND, A., PEDERSEN, N. C., HOFMANN-LEHMANN, R. & LUTZ, H. 2005. Feline coronavirus serotypes 1 and 2: seroprevalence and association with disease in Switzerland. *Clin Diagn Lab Immunol,* 12**,** 1209-15.

LAHAR, N., LEI, N. Y., WANG, J., JABAJI, Z., TUNG, S. C., JOSHI, V., LEWIS, M., STELZNER, M., MARTIN, M. G. & DUNN, J. C. 2011. Intestinal subepithelial myofibroblasts support in vitro and in vivo growth of human small intestinal epithelium. *PLoS One,* 6**,** e26898.

LAVER, T., HARRISON, J., O'NEILL, P. A., MOORE, K., FARBOS, A., PASZKIEWICZ, K. & STUDHOLME, D. J. 2015. Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomol Detect Quantif,* 3**,** 1-8.

LE PODER, S. 2011. Feline and canine coronaviruses: common genetic and pathobiological features. *Adv Virol,* 2011**,** 609465.

LEGENDRE, A. M. & BARTGES, J. W. 2009. Effect of Polyprenyl Immunostimulant on the survival times of three cats with the dry form of feline infectious peritonitis. *J Feline Med Surg,* 11**,** 624-6.

LEGENDRE, A. M., KURITZ, T., GALYON, G., BAYLOR, V. M. & HEIDEL, R. E. 2017. Polyprenyl Immunostimulant Treatment of Cats with Presumptive Non-Effusive Feline Infectious Peritonitis In a Field Study. *Front Vet Sci,* 4**,** 7.

LEWIS, C. S., PORTER, E., MATTHEWS, D., KIPAR, A., TASKER, S., HELPS, C. R. & SIDDELL, S. G. 2015. Genotyping coronaviruses associated with feline infectious peritonitis. *J Gen Virol,* 96**,** 1358-68.

LI, C., LIU, Q., KONG, F., GUO, D., ZHAI, J., SU, M. & SUN, D. 2018. Circulation and Genetic Diversity of Feline Coronavirus Type I and II From Clinically Healthy and FIP-Suspected Cats in China. *Transbound Emerg Dis*.

LI, F. 2015. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. *J Virol,* 89**,** 1954-64.

LI, F. 2016. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu Rev Virol*.

LICITRA, B. N., MILLET, J. K., REGAN, A. D., HAMILTON, B. S., RINALDI, V. D., DUHAMEL, G. E. & WHITTAKER, G. R. 2013. Mutation in spike protein cleavage site and pathogenesis of feline coronavirus. *Emerg Infect Dis,* 19**,** 1066-73.

LIN, C. N., SU, B. L., WANG, C. H., HSIEH, M. W., CHUEH, T. J. & CHUEH, L. L. 2009. Genetic diversity and correlation with feline infectious peritonitis of feline coronavirus type I and II: a 5-year study in Taiwan. *Vet Microbiol,* 136**,** 233-9.

LITSTER, A. L., POGRANICHNIY, R. & LIN, T. L. 2013. Diagnostic utility of a direct immunofluorescence test to detect feline coronavirus antigen in macrophages in effusive feline infectious peritonitis. *Vet J,* 198**,** 362-6.

LONGSTAFF, L., PORTER, E., CROSSLEY, V. J., HAYHOW, S. E., HELPS, C. R. & TASKER, S. 2015. Feline coronavirus quantitative reverse transcriptase polymerase chain reaction on effusion samples in cats with and without feline infectious peritonitis. *J Feline Med Surg*.

MAHE, M. M., AIHARA, E., SCHUMACHER, M. A., ZAVROS, Y., MONTROSE, M. H., HELMRATH, M. A., SATO, T. & SHROYER, N. F. 2013. Establishment of Gastrointestinal Epithelial Organoids. *Curr Protoc Mouse Biol,* 3**,** 217-40.

MALBON, A. J., MELI, M. L., BARKER, E. N., DAVIDSON, A. D., TASKER, S. & KIPAR, A. 2019. Inflammatory Mediators in the Mesenteric Lymph Nodes, Site of a Possible Intermediate Phase in the Immune Response to Feline Coronavirus and the Pathogenesis of Feline Infectious Peritonitis? *J Comp Pathol,* 166**,** 69-86.

MAYER, M. P. 2013. Hsp70 chaperone dynamics and molecular mechanism. *Trends Biochem Sci,* 38**,** 507-14.

MCBRIDE, R., VAN ZYL, M. & FIELDING, B. C. 2014. The coronavirus nucleocapsid is a multifunctional protein. *Viruses,* 6**,** 2991-3018.

MEAZZI, S., STRANIERI, A., LAUZI, S., BONSEMBIANTE, F., FERRO, S., PALTRINIERI, S. & GIORDANO, A. 2019. Feline gut microbiota composition in association with feline coronavirus infection: A pilot study. *Res Vet Sci,* 125**,** 272-278.

MELI, M., KIPAR, A., MULLER, C., JENAL, K., GONCZI, E., BOREL, N., GUNN-MOORE, D., CHALMERS, S., LIN, F., REINACHER, M. & LUTZ, H. 2004. High viral loads despite absence of clinical and pathological findings in cats experimentally infected with feline coronavirus (FCoV) type I and in naturally FCoV-infected cats. *J Feline Med Surg,* 6**,** 69-81.

MESZAROS, I., OLASZ, F., KADAR-HURKECZ, E., BALINT, A., HORNYAK, A., BELAK, S. & ZADORI, Z. 2018. Cellular localisation of the proteins of region 3 of feline enteric coronavirus. *Acta Vet Hung,* 66**,** 493-508.

MIGUEL, B., PHARR, G. T. & WANG, C. 2002. The role of feline aminopeptidase N as a receptor for infectious bronchitis virus. Brief review. *Arch Virol,* 147**,** 2047-56.

MIYOSHI, H. & STAPPENBECK, T. S. 2013. In vitro expansion and genetic modification of gastrointestinal stem cells in spheroid culture. *Nature Protocols,* 8**,** 2471-2482.

MURPHY, B. G., PERRON, M., MURAKAMI, E., BAUER, K., PARK, Y., ECKSTRAND, C., LIEPNIEKS, M. & PEDERSEN, N. C. 2018. The nucleoside analog GS-441524 strongly inhibits feline infectious peritonitis (FIP) virus in tissue culture and experimental cat infection studies. *Vet Microbiol,* 219**,** 226-233.

MURPHY, K., TRAVERS, P., WALPORT, M. & JANEWAY, C. 2012. *Janeway's immunobiology,* London, Garland Science ; London : Taylor & Francis [distributor].

MUSTAFFA-KAMAL, F., LIU, H., PEDERSEN, N. C. & SPARGER, E. E. 2019. Characterization of antiviral T cell responses during primary and secondary challenge of laboratory cats with feline infectious peritonitis virus (FIPV). *BMC Vet Res,* 15**,** 165.

NAMY, O., MORAN, S. J., STUART, D. I., GILBERT, R. J. & BRIERLEY, I. 2006. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature,* 441**,** 244-7.

NEUMAN, B. W., CHAMBERLAIN, P., BOWDEN, F. & JOSEPH, J. 2014. Atlas of coronavirus replicase structure. *Virus Res,* 194**,** 49-66.

NEUMAN, B. W., KISS, G., KUNDING, A. H., BHELLA, D., BAKSH, M. F., CONNELLY, S., DROESE, B., KLAUS, J. P., MAKINO, S., SAWICKI, S. G., SIDDELL, S. G., STAMOU, D. G., WILSON, I. A., KUHN, P. & BUCHMEIER, M. J. 2011. A structural analysis of M protein in coronavirus assembly and morphology. *J Struct Biol,* 174**,** 11-22.

NORDONE, S. K., STEVENS, R., LAVOY, A. S. & DEAN, G. A. 2005. Feline cytokine ELISPOT: issues in assay development. *Methods Mol Biol,* 302**,** 167-78.

NORRIS, J. M., BOSWARD, K. L., WHITE, J. D., BARAL, R. M., CATT, M. J. & MALIK, R. 2005. Clinicopathological findings associated with feline infectious peritonitis in Sydney, Australia: 42 cases (1990-2002). *Aust Vet J,* 83**,** 666-73.

OLSEN, C. W. 1993. A review of feline infectious peritonitis virus: molecular biology, immunopathogenesis, clinical aspects, and vaccination. *Vet Microbiol,* 36**,** 1-37.

OLSEN, C. W., CORAPI, W. V., NGICHABE, C. K., BAINES, J. D. & SCOTT, F. W. 1992. Monoclonal antibodies to the spike protein of feline infectious peritonitis virus mediate antibody-dependent enhancement of infection of feline macrophages. *J Virol,* 66**,** 956-65.

OLYSLAEGERS, D. A., DEDEURWAERDER, A., DESMARETS, L. M., VERMEULEN, B. L., DEWERCHIN, H. L. & NAUWYNCK, H. J. 2013. Altered expression of adhesion molecules on peripheral blood leukocytes in feline infectious peritonitis. *Vet Microbiol,* 166**,** 438-49.

ORTEGO, J., ESCORS, D., LAUDE, H. & ENJUANES, L. 2002. Generation of a replication-competent, propagation-deficient virus vector based on the transmissible gastroenteritis coronavirus genome. *J Virol,* 76**,** 11518-29.

OSUMI, T., MITSUI, I., LEUTENEGGER, C. M., OKABE, R., IDE, K. & NISHIFUJI, K. 2018. First identification of a single amino acid change in the spike protein region of feline coronavirus detected from a coronavirus-associated cutaneous nodule in a cat. *JFMS Open Rep,* 4**,** 2055116918801385.

OTSUKI, K., MAEDA, J., YAMAMOTO, H. & TSUBOKURA, M. 1979. Studies on avian infectious bronchitis virus (IBV). III. Interferon induction by and sensitivity to interferon of IBV. *Arch Virol,* 60**,** 249-55.

PALTRINIERI, S., GIORDANO, A., TRANQUILLO, V. & GUAZZETTI, S. 2007. Critical assessment of the diagnostic value of feline alpha1-acid glycoprotein for feline infectious peritonitis using the likelihood ratios approach. *J Vet Diagn Invest,* 19**,** 266-72.

PASSLACK, N., KOHN, B., DOHERR, M. G. & ZENTEK, J. 2017. Impact of Dietary Protein Concentration and Quality on Immune Function of Cats. *PLoS One,* 12**,** e0169822.

PASTERNAK, A. O., SPAAN, W. J. & SNIJDER, E. J. 2006. Nidovirus transcription: how to make sense...? *J Gen Virol,* 87**,** 1403-21.

PEDERSEN, N. C. 1987. Virologic and immunologic aspects of feline infectious peritonitis virus infection. *Adv Exp Med Biol,* 218**,** 529-50.

PEDERSEN, N. C. 2009. A review of feline infectious peritonitis virus infection: 1963-2008. *J Feline Med Surg,* 11**,** 225-58.

PEDERSEN, N. C. 2014a. An update on feline infectious peritonitis: diagnostics and therapeutics. *Vet J,* 201**,** 133-41.

PEDERSEN, N. C. 2014b. An update on feline infectious peritonitis: virology and immunopathogenesis. *Vet J,* 201**,** 123-32.

PEDERSEN, N. C., ALLEN, C. E. & LYONS, L. A. 2008. Pathogenesis of feline enteric coronavirus infection. *J Feline Med Surg,* 10**,** 529-41.
PEDERSEN, N. C., ECKSTRAND, C., LIU, H., LEUTENEGGER, C. & MURPHY, B. 2015. Levels of feline infectious peritonitis virus in blood, effusions, and various tissues and the role of lymphopenia in disease outcome following experimental infection. *Vet Microbiol,* 175**,** 157-66.
PEDERSEN, N. C., KIM, Y., LIU, H., GALASITI KANKANAMALAGE, A. C., ECKSTRAND, C., GROUTAS, W. C., BANNASCH, M., MEADOWS, J. M. & CHANG, K. O. 2017. Efficacy of a 3C-like protease inhibitor in treating various forms of acquired feline infectious peritonitis. *J Feline Med Surg***,** 1098612X17729626.
PEDERSEN, N. C., LIU, H., DURDEN, M. & LYONS, L. A. 2016. Natural resistance to experimental feline infectious peritonitis virus infection is decreased rather than increased by positive genetic selection. *Vet Immunol Immunopathol,* 171**,** 17-20.
PEDERSEN, N. C., LIU, H., GANDOLFI, B. & LYONS, L. A. 2014. The influence of age and genetics on natural resistance to experimentally induced feline infectious peritonitis. *Vet Immunol Immunopathol,* 162**,** 33-40.
PEDERSEN, N. C., LIU, H., SCARLETT, J., LEUTENEGGER, C. M., GOLOVKO, L., KENNEDY, H. & KAMAL, F. M. 2012. Feline infectious peritonitis: role of the feline coronavirus 3c gene in intestinal tropism and pathogenicity based upon isolates from resident and adopted shelter cats. *Virus Res,* 165**,** 17-28.
PEDERSEN, N. C., PERRON, M., BANNASCH, M., MONTGOMERY, E., MURAKAMI, E., LIEPNIEKS, M. & LIU, H. 2019. Efficacy and safety of the nucleoside analog GS-441524 for treatment of cats with naturally occurring feline infectious peritonitis. *J Feline Med Surg***,** 1098612X19825701.
PEREZ-VARGAS, J., ROMERO, P., LOPEZ, S. & ARIAS, C. F. 2006. The peptide-binding and ATPase domains of recombinant hsc70 are required to interact with rotavirus and reduce its infectivity. *J Virol,* 80**,** 3322-31.
PESTEANU-SOMOGYI, L. D., RADZAI, C. & PRESSLER, B. M. 2006. Prevalence of feline infectious peritonitis in specific cat breeds. *J Feline Med Surg,* 8**,** 1-5.
PONCELET, L., COPPENS, A., PEETERS, D., BIANCHI, E., GRANT, C. K. & KADHIM, H. 2008. Detection of antigenic heterogeneity in feline coronavirus nucleocapsid in feline pyogranulomatous meningoencephalitis. *Vet Pathol,* 45**,** 140-53.
PORTER, E., TASKER, S., DAY, M. J., HARLEY, R., KIPAR, A., SIDDELL, S. G. & HELPS, C. R. 2014. Amino acid changes in the spike protein of feline coronavirus correlate with systemic spread of virus from the intestine and not with feline infectious peritonitis. *Vet Res,* 45**,** 49.
POWELL, R. H. & BEHNKE, M. S. 2017. WRN conditioned media is sufficient for in vitro propagation of intestinal organoids from large farm and small companion animals. *Biology Open,* 6**,** 698-705.
PUJHARI, S., BRUSTOLIN, M., MACIAS, V. M., NISSLY, R. H., NOMURA, M., KUCHIPUDI, S. V. & RASGON, J. L. 2019. Heat shock protein 70 (Hsp70) mediates Zika virus entry, replication, and egress from host cells. *Emerg Microbes Infect,* 8**,** 8-16.
QUIRKE, G. 2016. *Coronavirus Entry Mechanisms.* MRes, University of Bristol.
R&D. 2016. *Troubleshooting guide: ELISpot* [Online]. Available: https://www.rndsystems.com/resources/technical/troubleshooting-guide-elispot [Accessed 25 February 2016].
RAJ, V. S., MOU, H., SMITS, S. L., DEKKERS, D. H., MULLER, M. A., DIJKMAN, R., MUTH, D., DEMMERS, J. A., ZAKI, A., FOUCHIER, R. A., THIEL, V., DROSTEN, C., ROTTIER, P. J., OSTERHAUS, A. D., BOSCH, B. J. & HAAGMANS, B. L. 2013. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature,* 495**,** 251-4.
RAMAN, S. & BRIAN, D. A. 2005. Stem-loop IV in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. *J Virol,* 79**,** 12434-46.

RAMANI, S., CRAWFORD, S. E., BLUTT, S. E. & ESTES, M. K. 2018. Human organoid cultures: transformative new tools for human virus studies. *Curr Opin Virol,* 29**,** 79-86.
REDFORD, T. & AL-DISSI, A. N. 2019. Feline infectious peritonitis in a cat presented because of papular skin lesions. *Can Vet J,* 60**,** 183-185.
REGAN, A. D., OUSTEROUT, D. G. & WHITTAKER, G. R. 2010. Feline lectin activity is critical for the cellular entry of feline infectious peritonitis virus. *J Virol,* 84**,** 7917-21.
REGAN, A. D., SHRAYBMAN, R., COHEN, R. D. & WHITTAKER, G. R. 2008. Differential role for low pH and cathepsin-mediated cleavage of the viral spike protein during entry of serotype II feline coronaviruses. *Vet Microbiol,* 132**,** 235-48.
REGAN, A. D. & WHITTAKER, G. R. 2008. Utilization of DC-SIGN for entry of feline coronaviruses into host cells. *J Virol,* 82**,** 11992-6.
REYES-DEL VALLE, J., CHAVEZ-SALINAS, S., MEDINA, F. & DEL ANGEL, R. M. 2005. Heat shock protein 90 and heat shock protein 70 are components of dengue virus receptor complex in human cells. *J Virol,* 79**,** 4557-67.
RIEMER, F., KUEHNER, K. A., RITZ, S., SAUTER-LOUIS, C. & HARTMANN, K. 2015. Clinical and laboratory features of cats with feline infectious peritonitis - a retrospective study of 231 confirmed cases (2000-2010). *J Feline Med Surg*.
RITZ, S., EGBERINK, H. & HARTMANN, K. 2007. Effect of feline interferon-omega on the survival time and quality of life of cats with feline infectious peritonitis. *J Vet Intern Med,* 21**,** 1193-7.
ROBERTS, J. D., THAPALIYA, A., MARTINEZ-LUMBRERAS, S., KRYSZTOFINSKA, E. M. & ISAACSON, R. L. 2015. Structural and Functional Insights into Small, Glutamine-Rich, Tetratricopeptide Repeat Protein Alpha. *Front Mol Biosci,* 2**,** 71.
ROHRBACH, B. W., LEGENDRE, A. M., BALDWIN, C. A., LEIN, D. H., REED, W. M. & WILSON, R. B. 2001. Epidemiology of feline infectious peritonitis among cats examined at veterinary medical teaching hospitals. *J Am Vet Med Assoc,* 218**,** 1111-5.
ROTTIER, P. J. 1999. The molecular dynamics of feline coronaviruses. *Vet Microbiol,* 69**,** 117-25.
ROTTIER, P. J., NAKAMURA, K., SCHELLEN, P., VOLDERS, H. & HAIJEMA, B. J. 2005. Acquisition of macrophage tropism during the pathogenesis of feline infectious peritonitis is determined by mutations in the feline coronavirus spike protein. *J Virol,* 79**,** 14122-30.
SABSHIN, S. J., LEVY, J. K., TUPLER, T., TUCKER, S. J., GREINER, E. C. & LEUTENEGGER, C. M. 2012. Enteropathogens identified in cats entering a Florida animal shelter with normal feces or diarrhea. *J Am Vet Med Assoc,* 241**,** 331-7.
SANGL, L., MATIASEK, K., FELTEN, S., GRUNDL, S., BERGMANN, M., BALZER, H. J., PANTCHEV, N., LEUTENEGGER, C. M. & HARTMANN, K. 2018. Detection of feline coronavirus mutations in paraffin-embedded tissues in cats with feline infectious peritonitis and controls. *J Feline Med Surg***,** 1098612X18762883.
SATO, T., STANGE, D. E., FERRANTE, M., VRIES, R. G., VAN ES, J. H., VAN DEN BRINK, S., VAN HOUDT, W. J., PRONK, A., VAN GORP, J., SIERSEMA, P. D. & CLEVERS, H. 2011. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology,* 141**,** 1762-72.
SATO, T., VRIES, R. G., SNIPPERT, H. J., VAN DE WETERING, M., BARKER, N., STANGE, D. E., VAN ES, J. H., ABO, A., KUJALA, P., PETERS, P. J. & CLEVERS, H. 2009. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature,* 459**,** 262-5.
SATOH, R., FURUKAWA, T., KOTAKE, M., TAKANO, T., MOTOKAWA, K., GEMMA, T., WATANABE, R., ARAI, S. & HOHDATSU, T. 2011a. Screening and identification of T helper 1 and linear immunodominant antibody-binding epitopes in the spike 2 domain and the nucleocapsid protein of feline infectious peritonitis virus. *Vaccine,* 29**,** 1791-800.
SATOH, R., KAKU, A., SATOMURA, M., KOHORI, M., NOURA, K., FURUKAWA, T., KOTAKE, M., TAKANO, T. & HOHDATSU, T. 2011b. Development of monoclonal antibodies (MAbs) to feline interferon (fIFN)-gamma as tools to evaluate cellular immune responses to feline infectious peritonitis virus (FIPV). *J Feline Med Surg,* 13**,** 427-35.

SATOH, R., KOBAYASHI, H., TAKANO, T., MOTOKAWA, K., KUSUHARA, H. & HOHDATSU, T. 2010. Characterization of T helper (Th)1- and Th2-type immune responses caused by baculovirus-expressed protein derived from the S2 domain of feline infectious peritonitis virus, and exploration of the Th1 and Th2 epitopes in a mouse model. *Microbiol Immunol,* 54**,** 726-33.

SAWICKI, S. G., SAWICKI, D. L. & SIDDELL, S. G. 2007. A contemporary view of coronavirus transcription. *J Virol,* 81**,** 20-9.

SAXENA, K., BLUTT, S. E., ETTAYEBI, K., ZENG, X. L., BROUGHMAN, J. R., CRAWFORD, S. E., KARANDIKAR, U. C., SASTRI, N. P., CONNER, M. E., OPEKUN, A. R., GRAHAM, D. Y., QURESHI, W., SHERMAN, V., FOULKE-ABEL, J., IN, J., KOVBASNJUK, O., ZACHOS, N. C., DONOWITZ, M. & ESTES, M. K. 2016. Human Intestinal Enteroids: a New Model To Study Human Rotavirus Infection, Host Restriction, and Pathophysiology. *J Virol,* 90**,** 43-56.

SCHMITTEL, A., KEILHOLZ, U., BAUER, S., KUHNE, U., STEVANOVIC, S., THIEL, E. & SCHEIBENBOGEN, C. 2001. Application of the IFN-gamma ELISPOT assay to quantify T cell responses against proteins. *J Immunol Methods,* 247**,** 17-24.

SCHOEMAN, D. & FIELDING, B. C. 2019. Coronavirus envelope protein: current knowledge. *Virol J,* 16**,** 69.

SCHWEGMANN-WESSELS, C. & HERRLER, G. 2006. Sialic acids as receptor determinants for coronaviruses. *Glycoconj J,* 23**,** 51-8.

SCOBEY, T., YOUNT, B. L., SIMS, A. C., DONALDSON, E. F., AGNIHOTHRAM, S. S., MENACHERY, V. D., GRAHAM, R. L., SWANSTROM, J., BOVE, P. F., KIM, J. D., GREGO, S., RANDELL, S. H. & BARIC, R. S. 2013. Reverse genetics with a full-length infectious cDNA of the Middle East respiratory syndrome coronavirus. *Proc Natl Acad Sci U S A,* 110**,** 16157-62.

SHARIFI TABAR, M., HESARAKI, M., ESFANDIARI, F., SAHRANESHIN SAMANI, F., VAKILIAN, H. & BAHARVAND, H. 2015. Evaluating Electroporation and Lipofectamine Approaches for Transient and Stable Transgene Expressions in Human Fibroblasts and Embryonic Stem Cells. *Cell J,* 17**,** 438-50.

SHIBA, N., MAEDA, K., KATO, H., MOCHIZUKI, M. & IWATA, H. 2007. Differentiation of feline coronavirus type I and II infections by virus neutralization test. *Vet Microbiol,* 124**,** 348-52.

SHIRATO, K., CHANG, H. W. & ROTTIER, P. J. M. 2018. Differential susceptibility of macrophages to serotype II feline coronaviruses correlates with differences in the viral spike protein. *Virus Res*.

SIDDELL, S. 1995. *The coronaviridae,* New York ; London, Plenum Press.

SLOTA, M., LIM, J. B., DANG, Y. & DISIS, M. L. 2011. ELISpot for measuring human immune responses to vaccines. *Expert Rev Vaccines,* 10**,** 299-306.

SOLA, I., ALMAZAN, F., ZUNIGA, S. & ENJUANES, L. 2015. Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annu Rev Virol,* 2**,** 265-88.

SOMA, T., SAITO, N., KAWAGUCHI, M. & SASAI, K. 2018. Feline coronavirus antibody titer in cerebrospinal fluid from cats with neurological signs. *J Vet Med Sci,* 80**,** 59-62.

SPARKES, A. H., GRUFFYDD-JONES, T. J. & HARBOUR, D. A. 1991. Feline infectious peritonitis: a review of clinicopathological changes in 65 cases, and a critical assessment of their diagnostic value. *Vet Rec,* 129**,** 209-12.

STOBART, C. C. & MOORE, M. L. 2014. RNA virus reverse genetics and vaccine design. *Viruses,* 6**,** 2531-50.

STOHLMAN, S. A., BARIC, R. S., NELSON, G. N., SOE, L. H., WELTER, L. M. & DEANS, R. J. 1988. Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J Virol,* 62**,** 4288-95.

STRANIERI, A., GIORDANO, A., PALTRINIERI, S., GIUDICE, C., CANNITO, V. & LAUZI, S. 2018. Comparison of the performance of laboratory tests in the diagnosis of feline infectious peritonitis. *J Vet Diagn Invest***,** 1040638718756460.

SZCZEPANSKI, A., OWCZAREK, K., BZOWSKA, M., GULA, K., DREBOT, I., OCHMAN, M., MAKSYM, B., RAJFUR, Z., MITCHELL, J. A. & PYRC, K. 2019. Canine Respiratory Coronavirus, Bovine Coronavirus, and Human Coronavirus OC43: Receptors and Attachment Factors. *Viruses,* 11.

TAKANO, T., AKIYAMA, M., DOKI, T. & HOHDATSU, T. 2019a. Antiviral activity of itraconazole against type I feline coronavirus infection. *Vet Res,* 50**,** 5.
TAKANO, T., AZUMA, N., SATOH, M., TODA, A., HASHIDA, Y., SATOH, R. & HOHDATSU, T. 2009. Neutrophil survival factors (TNF-alpha, GM-CSF, and G-CSF) produced by macrophages in cats infected with feline infectious peritonitis virus contribute to the pathogenesis of granulomatous lesions. *Arch Virol,* 154**,** 775-81.
TAKANO, T., HOHDATSU, T., HASHIDA, Y., KANEKO, Y., TANABE, M. & KOYAMA, H. 2007. A "possible" involvement of TNF-alpha in apoptosis induction in peripheral blood lymphocytes of cats with feline infectious peritonitis. *Vet Microbiol,* 119**,** 121-31.
TAKANO, T., KATADA, Y., MORITOH, S., OGASAWARA, M., SATOH, K., SATOH, R., TANABE, M. & HOHDATSU, T. 2008a. Analysis of the mechanism of antibody-dependent enhancement of feline infectious peritonitis virus infection: aminopeptidase N is not important and a process of acidification of the endosome is necessary. *J Gen Virol,* 89**,** 1025-9.
TAKANO, T., KATOH, Y., DOKI, T. & HOHDATSU, T. 2013. Effect of chloroquine on feline infectious peritonitis virus infection in vitro and in vivo. *Antiviral Res,* 99**,** 100-7.
TAKANO, T., KAWAKAMI, C., YAMADA, S., SATOH, R. & HOHDATSU, T. 2008b. Antibody-dependent enhancement occurs upon re-infection with the identical serotype virus in feline infectious peritonitis virus infection. *J Vet Med Sci,* 70**,** 1315-21.
TAKANO, T., MORIOKA, H., GOMI, K., TOMIZAWA, K., DOKI, T. & HOHDATSU, T. 2014. Screening and identification of T helper 1 and linear immunodominant antibody-binding epitopes in spike 1 domain and membrane protein of feline infectious peritonitis virus. *Vaccine,* 32**,** 1834-40.
TAKANO, T., YAMADA, S., DOKI, T. & HOHDATSU, T. 2019b. Pathogenesis of oral type I feline infectious peritonitis virus (FIPV) infection: antibody-dependent enhancement infection of cats with type I FIPV via the oral route. *J Vet Med Sci*.
TANAKA, Y., SATO, Y. & SASAKI, T. 2013. Suppression of coronavirus replication by cyclophilin inhibitors. *Viruses,* 5**,** 1250-60.
TASKER, S. 2018. Diagnosis of feline infectious peritonitis: Update on evidence supporting available tests. *J Feline Med Surg,* 20**,** 228-243.
TEKES, G., HOFMANN-LEHMANN, R., BANK-WOLF, B., MAIER, R., THIEL, H. J. & THIEL, V. 2010. Chimeric feline coronaviruses that encode type II spike protein on type I genetic background display accelerated viral growth and altered receptor usage. *J Virol,* 84**,** 1326-33.
TEKES, G., HOFMANN-LEHMANN, R., STALLKAMP, I., THIEL, V. & THIEL, H. J. 2008. Genome organization and reverse genetic analysis of a type I feline coronavirus. *J Virol,* 82**,** 1851-9.
TEKES, G., SPIES, D., BANK-WOLF, B., THIEL, V. & THIEL, H. J. 2012. A reverse genetics approach to study feline infectious peritonitis. *J Virol,* 86**,** 6994-8.
THERMO-FISHER. 2015. *Transfection Reagent FAQs* [Online]. Available: https://www.thermofisher.com/uk/en/home/life-science/cell-culture/transfection/transfection-support/transfection-reagent-faqs.html#faq2 [Accessed 28 July 2016].
THIEL, V., HEROLD, J., SCHELLE, B. & SIDDELL, S. G. 2001. Viral replicase gene products suffice for coronavirus discontinuous transcription. *J Virol,* 75**,** 6676-81.
THIEL, V., THIEL, H. J. & TEKES, G. 2014. Tackling feline infectious peritonitis via reverse genetics. *Bioengineered,* 5**,** 396-400.
TIAN, J., HU, X., LIU, D., WU, H. & QU, L. 2016. Identification of Inonotus obliquus polysaccharide with broad-spectrum antiviral activity against multi-feline viruses. *Int J Biol Macromol*.
TRESNAN, D. B., LEVIS, R. & HOLMES, K. V. 1996. Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I. *J Virol,* 70**,** 8669-74.
UCHIDA, R., SAITO, Y., NOGAMI, K., KAJIYAMA, Y., SUZUKI, Y., KAWASE, Y., NAKAOKA, T., MURAMATSU, T., KIMURA, M. & SAITO, H. 2019. Epigenetic silencing of Lgr5 induces senescence of intestinal epithelial organoids during the process of aging. *NPJ Aging Mech Dis,* 5**,** 1.

VAN DEN WORM, S. H., ERIKSSON, K. K., ZEVENHOVEN, J. C., WEBER, F., ZUST, R., KURI, T., DIJKMAN, R., CHANG, G., SIDDELL, S. G., SNIJDER, E. J., THIEL, V. & DAVIDSON, A. D. 2012. Reverse genetics of SARS-related coronavirus using vaccinia virus-based recombination. *PLoS One,* 7**,** e32857.

VAN DER HEE, B., LOONEN, L. M. P., TAVERNE, N., TAVERNE-THIELE, J. J., SMIDT, H. & WELLS, J. M. 2018. Optimized procedures for generating an enhanced, near physiological 2D culture system from porcine intestinal organoids. *Stem Cell Res,* 28**,** 165-171.

VAN HAMME, E., DESMARETS, L., DEWERCHIN, H. L. & NAUWYNCK, H. J. 2011. Intriguing interplay between feline infectious peritonitis virus and its receptors during entry in primary feline monocytes. *Virus Res,* 160**,** 32-9.

VAN HAMME, E., DEWERCHIN, H. L., CORNELISSEN, E. & NAUWYNCK, H. J. 2007. Attachment and internalization of feline infectious peritonitis virus in feline blood monocytes and Crandell feline kidney cells. *J Gen Virol,* 88**,** 2527-32.

VENNEMA, H., DE GROOT, R. J., HARBOUR, D. A., DALDERUP, M., GRUFFYDD-JONES, T., HORZINEK, M. C. & SPAAN, W. J. 1990. Early death after feline infectious peritonitis virus challenge due to recombinant vaccinia virus immunization. *J Virol,* 64**,** 1407-9.

VENNEMA, H., HEIJNEN, L., ROTTIER, P. J., HORZINEK, M. C. & SPAAN, W. J. 1992. A novel glycoprotein of feline infectious peritonitis coronavirus contains a KDEL-like endoplasmic reticulum retention signal. *J Virol,* 66**,** 4951-6.

VENNEMA, H., POLAND, A., FOLEY, J. & PEDERSEN, N. C. 1998. Feline infectious peritonitis viruses arise by mutation from endemic feline enteric coronaviruses. *Virology,* 243**,** 150-7.

VERMEULEN, B. L., DEVRIENDT, B., OLYSLAEGERS, D. A., DEDEURWAERDER, A., DESMARETS, L. M., FAVOREEL, H. W., DEWERCHIN, H. L. & NAUWYNCK, H. J. 2013. Suppression of NK cells and regulatory T lymphocytes in cats naturally infected with feline infectious peritonitis virus. *Vet Microbiol,* 164**,** 46-59.

VIEHWEGER, A., KRAUTWURST, S., LAMKIEWICZ, K., MADHUGIRI, R., ZIEBUHR, J., HOLZER, M. & MARZ, M. 2019. Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. *Genome Res,* 29**,** 1545-1554.

WANG, Y. T., HSIEH, L. E., DAI, Y. R. & CHUEH, L. L. 2014. Polymorphisms in the feline TNFA and CD209 genes are associated with the outcome of feline coronavirus infection. *Vet Res,* 45**,** 123.

WATANABE, R., ECKSTRAND, C., LIU, H. & PEDERSEN, N. C. 2018. Characterization of peritoneal cells from cats with experimentally-induced feline infectious peritonitis (FIP) using RNA-seq. *Vet Res,* 49**,** 81.

WATARI, T., KANESHIMA, T., TSUJIMOTO, H., ONO, K. & HASEGAWA, A. 1998. Effect of thromboxane synthetase inhibitor on feline infectious peritonitis in cats. *J Vet Med Sci,* 60**,** 657-9.

WEISS, R. C., COX, N. R. & MARTINEZ, M. L. 1993. Evaluation of free or liposome-encapsulated ribavirin for antiviral therapy of experimentally induced feline infectious peritonitis. *Res Vet Sci,* 55**,** 162-72.

WEISS, R. C., COX, N. R. & OOSTROM-RAM, T. 1990. Effect of interferon or Propionibacterium acnes on the course of experimentally induced feline infectious peritonitis in specific-pathogen-free and random-source cats. *Am J Vet Res,* 51**,** 726-33.

WIDAGDO, W., OKBA, N. M. A., LI, W., DE JONG, A., DE SWART, R. L., BEGEMAN, L., VAN DEN BRAND, J. M. A., BOSCH, B. J. & HAAGMANS, B. L. 2019. Species-Specific Colocalization of Middle East Respiratory Syndrome Coronavirus Attachment and Entry Receptors. *J Virol,* 93.

WILLIAMS, R. K., JIANG, G. S. & HOLMES, K. V. 1991. Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. *Proc Natl Acad Sci U S A,* 88**,** 5533-6.

WU, K., LI, W., PENG, G. & LI, F. 2009. Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. *Proc Natl Acad Sci U S A,* 106**,** 19970-4.

YEAGER, C. L., ASHMUN, R. A., WILLIAMS, R. K., CARDELLICHIO, C. B., SHAPIRO, L. H., LOOK, A. T. & HOLMES, K. V. 1992. Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature,* 357**,** 420-2.

YIP, C. W., HON, C. C., SHI, M., LAM, T. T., CHOW, K. Y., ZENG, F. & LEUNG, F. C. 2009. Phylogenetic perspectives on the epidemiology and origins of SARS and SARS-like coronaviruses. *Infect Genet Evol,* 9**,** 1185-96.

YOUNT, B., CURTIS, K. M. & BARIC, R. S. 2000. Strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. *J Virol,* 74**,** 10600-11.

ZHANG, Z., YANG, X., XU, P., WU, X., ZHOU, L. & WANG, H. 2017. Heat shock protein 70 in lung and kidney of specific-pathogen-free chickens is a receptor-associated protein that interacts with the binding domain of the spike protein of infectious bronchitis virus. *Arch Virol,* 162**,** 1625-1631.

ZHU, Y. Z., CAO, M. M., WANG, W. B., WANG, W., REN, H., ZHAO, P. & QI, Z. T. 2012. Association of heat-shock protein 70 with lipid rafts is required for Japanese encephalitis virus infection in Huh7 cells. *J Gen Virol,* 93**,** 61-71.

ZIMMER, G. 2010. RNA replicons - a new approach for influenza virus immunoprophylaxis. *Viruses,* 2**,** 413-34.

Bait protein coding sequences

NNNN FCoV S1
NNNN Human IgG Fc
NNNN Restriction site

Type 1 bait protein

GCGGCCGCAGGTACCACCATGATCCTGCTGCTGGTGCTGTTCAGCGTGGTGGGCGCCCACGACGCCCCCCACGGC
GTGACCCTGCCCCAGTTCAACACCAGCTACAACAACGACAAGTTCGAGCTGAACTTCTACAACTTCCTGCAGACC
TGGGACATCCCCCCCCAACACCGAGACCATCCTGGGCGGCTACCTGCCCTACTGCGGCACCGGCGCCAACTGCGGC
TGGTACAACTTCGTGTACCAGCAGAACGTGGGCAGCAACGGCAAGTACAGCTACATCAACACCCAGAACCTGAAC
ATCCCCAACGTGCACGGCGTGTACTTCGACGTGAGAGAGCACAACAGCGACGGCGTGTGGGACCTGAGAGACAGA
GTGGGCCTGCTGATCGCCATCCACGGCAAGAGCCAGTACAGCCTGCTGATGGTGCTGCAGGACAACGTGGAGGAG
AACCAGCCCCACGTGGCCGTGAAGATCTGCCACTGGAAGCCCGGCAACATCAGCAGCATCCACCAGTTCAGCGTG
AACCTGGGCGACGGCGGCCAGTGCGTGTTCAACCAGAGATTCAGCCTGGACACCATCCTGACCACCAACGACTTC
TACGGCTTCCAGTGGACCAACAACTACGTGAACATCTACCTGGGCGGCACCATCACCAAAGTGTGGGTGGAGAAC
GACTGGAGCGTGGTGGAGAGCAGCATCAGCTACCACTGGAGCAGAATCAACTACGGCTACTACATGCAGTTCGTG
AACAGAACCACCTACTACGTGTACAACAACACCGGCGGCGCCAACTACACCCACCTGCAGCTGGAGGAGTGCCAC
AGCGAGTACTGCGCCGGCTACGCCAAGAACGTGTTCGTGCCCATCGACGGCAAGATCCCCGACGGCTTCAGCTTC
AGCAACTGGTTCCTGCTGAGCGACAAGAGCACCCTGGTGCAGGGCAGAGTGCTGAGCAAGCAGCCCGTGTTCGTG
CAGTGCCTGAGAGCCGTGCCCAGCTGGAGCAACAACAGCGCCGTGGTGCACTTCAGCAACGACGACTTCTGCCCC
AACGTGACCGCCGAGGTGCTGAGATTCAACCTGAACTTCAGCGACACCGACGTGTACGTGGCCAGCAACAGCGAC
GACAGACTGTACTTCACCTTCGAGGACAACACCACCGCCGGCGTGGCCTGCTACAGCAGCGCCAACATCACCGAC
TACAAGCCCAACACCAACGCCAGCAGCCAGATCCCCTTCGGCAAGACCACCCACAGCTACTTCTGCTTCGCCAAC
TTCAGCAGCAGCCACGTGACCCAGTTCCTGGGCATCCTGCCCCCCACCGTGAGAGAGTTCGCCTTCGGCAAGGAC
GGCAGCATCTTCGTGAACGGCTACAAGTACTTCAGCTTCCCCCCCATCAAGAGCGTGAACTTCAGCATCAGCAGC
GTGGAGAACTTCGGCTTCTGGACCATCGCCTACACCAACTACACCGACGTGATGGTGGACGTGAACGGCACCGGC
ATCACCAGACTGTTCTACTGCGACAGCCCCATCAACAGAATCAAGTGCCAGCAGCTGAAGTACGAGCTGCCCGAC
GGCTTCTACAGCGCCAGCATGCTGGTGAAGAAGGACCTGCCCAAGACCTTCGTGACCATGCCCCAGTTCTACAAC
TGGATGAACGTGACCCTGCACGTGGTGCTGAACGACACCGAGAAGAAGGCCGACATCATCCTGGCCAAGGCCAGC
GAGCTGGCCAGCCTGGCCGACGTGCACTTCGAGATCGCCCAGGCCAACGGCAGCGTGGTGAACGCCACCAGCCTG
TGCGTGCAGACCAGACAGCTGGCCCTGTTCTACAAGTACACCAGCCTGCAGGGCCTGTACACCTACAGCAACCTG
GTGGAGCTGCAGAACTACGACTGCCCCTTCAGCCCCCACCAGTTCAACAACTACCTGCAGTTCGAGACCCTGTGC
TTCGACACCAGCCCCGCCGTGGCCGGCTGCAAGTGGAGCCTGGTGCACGACGTGAGATGGAGAACCCAGTTCGCC
ACCATCACCATCAGCTACAAGGACGGCGCCAAGATCACCACCATGCCCAAGGCCAAGCTGGGCTTCCAGGACATC
AGCAACATCGTGAAGGACGAGTGCACCGACTACAACATCTACGGCTTCCAGGGCACCGGCATCATCAGAAACACC
ACCACCAGAATCGTGGCCGGCCTGTACTACACCAGCATCAGCGGCGACCTGCTGGCCTTCAAGAACAGCACCACC
GGCGAGATCTTCACCGTGGTGCCCTGCGGCCTGACCGCCCAGGCCGCCGTGATCAACGACGAGATCGTGGGCGTG
ATCACCGCCGTGAACCAGACCGACCTGTTCGAGTTCGTGAACCACACCCAGGCCGACAAGACCCACACCTGCCCC
CCCTGCCCCGCCCCCGAGCTGCTGGGCGGCCCCAGCGTGTTCCTGTTCCCCCCCAAGCCCAAGGACACCCTGATG
ATCAGCAGAACCCCCGAGGTGACCTGCGTGGTGGTGGACGTGAGCCACGAGGACCCCGAGGTGAAGTTCAACTGG
TACGTGGACGGCGTGGAGGTGCACAACGCCAAGACCAAGCCCAGAGAGGAGCAGTACAACAGCACCTACAGAGTG
GTCAGCGTGCTGACCGTGCTGCACCAGGACTGGCTGAACGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAGGCC
CTGCCCGCCCCCATCGAGAAGACCATCAGCAAGGCCAAGGGCCAGCCCAGAGAGCCCCAGGTGTACACCCTGCCC
CCCAGCAGAGAGGAGATGACCAAGAACCAGGTCTCCCTGACCTGCCTGGTGAAGGGCTTCTACCCCAGCGACATC
GCCGTGGAGTGGGAGAGCAACGGCCAGCCCGAGAACAACTACAAGACCACCCCCCCCGTGCTGGACAGCGACGGC
AGCTTCTTCCTGTACAGCAAGCTGACCGTGGACAAGAGCAGATGGCAGCAGGGCAACGTGTTCAGCTGCAGCGTG
ATGCACGAGGCCCTGCACAACCACTACACCCAGAAGAGCCTGAGCCTGAGCCCCGGCAAGCTCGAGTTTAAAC

Type 2 bait protein

GCGGCCGCAGGTACCACCATGATCGTGCTGGTGACCTGCCTGCTGCTGCTGTGCAGCTACCACACCGTGCTGAGC
ACCACCAACAACGAGTGCATCCAGGTGAACGTGACCCAGCTGGCCGGCAACGAGAACCTGATCAGAGACTTCCTG
TTCAGCAACTTCAAGGAGGAGGGCAGCGTGGTGGTGGGCGGCTACTACCCCACCGAGGTGTGGTACAACTGCAGC
AGAACCGCCAGAACCACCGCCTTCCAGTACTTCAACAACATCCACGCCTTCTACTTCGTGATGGAGGCCATGGAG
AACAGCACCGGCAACGCCAGAGGCAAGCCCCTGCTGTTCCACGTGCACGGCGAGCCCGTGAGCGTGATCATCAGC
GCCTACAGAGACGACGTGCAGCAGAGACCCCTGCTGAAGCACGGCCTGGTGTGCATCACCAAGAACAGACACATC
AACTACGAGCAGTTCACCAGCAACCAGTGGAACAGCACCTGCACCGGCGCCGACAGAAAGATCCCCTTCAGCGTG
ATCCCCACCGACAACGGCACCAAGATCTACGGCCTGGAGTGGAACGACGACTTCGTGACCGCCTACATCAGCGGC
AGAAGCTACCACCTGAACATCAACACCAACTGGTTCAACAACGTGACCCTGCTGTACAGCAGAAGCAGCACCGCC
ACCTGGGAGTACAGCGCCGCCTACGCCTACCAGGGCGTGAGCAACTTCACCTACTACAAGCTGAACAACACCAAC
GGCCTGAAGACCTACGAGCTGTGCGAGGACTACGAGCACTGCACCGGCTACGCCACCAACGTGTTCGCCCCCACC
AGCGGCGGCTACATCCCCGACGGCTTCAGCTTCAACAACTGGTTCCTGCTGACCAACAGCAGCACCTTCGTGAGC
GGCAGATTCGTGACCAACCAGCCCCTGCTGATCAACTGCCTGTGGCCCGTGCCCAGCTTCGGCGTGGCCGCCCAG
GAGTTCTGCTTCGAGGGCGCCCAGTTCAGCCAGTGCAACGGCGTGAGCCTGAACAACACCGTGGACGTGATCAGA
TTCAACCTGAACTTCACCGCCGACGTGCAGAGCGGCATGGGCGCCACCGTGTTCAGCCTGAACACCACCGGCGGC
GTGATCCTGGAGATCAGCTGCTACAGCGACACCGTGAGCGAGAGCAGCAGCTACAGCTACGGCGAGATCCCCTTC
GGCATCACCGACGGCCCCAGATACTGCTACGTGCTGTACAACGGCACCGCCCTGAAGTACCTGGGCACCCTGCCC
CCCAGCGTGAAGGAGATCGCCATCAGCAAGTGGGGCCACTTCTACATCAACGGCTACAACTTCTTCAGCACCTTC
CCCATCGGCTGCATCAGCTTCAACCTGACCACCGGCGTGAGCGGCGCCTTCTGGACCATCGCCTACACCAGCTAC
ACCGAGGCCCTGGTGCAGGTGGAGAACACCGCCATCAAGAACGTGACCTACTGCAACAGCCACATCAACAACATC
AAGTGCAGCCAGCTGACCGCCAACCTGAACAACGGCTTCTACCCCGTGGCCAGCAGCGAGGTGGGCTTCGTGAAC
AAGAGCGTGGTGCTGCTGCCCAGCTTCTTCACCTACACCGCCGTGAACATCACCATCGACCTGGGCATGAAGCTG
AGCGGCTACGGCCAGCCCATCGCCAGCACCCTGAGCAACATCACCCTGCCCATGCAGGACAACAACACCGACGTG
TACTGCATCAGAAGCAACCAGTTCAGCGTGTACGTGCACAGCACCTGCAAGAGCAGCCTGTGGGACAACATCTTC
AACCAGGACTGCACCGACGTGCTGGAGGCCACCGCCGTGATCAAGACCGGCACCTGCCCCTTCAGCTTCGACAAG
CTGAACAACTACCTGACCTTCAACAAGTTCTGCCTGAGCCTGAGCCCCGTGGGCGCCAACTGCAAGTTCGACGTG
GCCGCCAGAACCAGAACCAACGAGCAGGTCGTGAGAAGCCTGTACGTGATCTACGAGGAGGGCGACAACATCGTG
GGCGTGCCCAGCGACAACAGCGGCCTGCACGACCTGAGCGTGCTGCACCTGGACAGCTGCACCGACTACAACATC
TACGGCAGAACCGGCGTGGGCATCATCAGAAGAACCAACAGCACCCTGCTGAGCGGCCTGTACTACACCAGCCTG
AGCGGCGACCTGCTGGGCTTCAAGAACGTGAGCGACGGCGTGATCTACAGCGTGACCCCCTGCGACGTGAGCGTG
CAGGCCGCCGTGATCGACGGCGCCATCGTGGGCGCCATGACCAGCATCAACAGCGAGCTGCTGGGCCTGACCCAC
TGGACCACCACCCCCAACTTCTACTACTACAGCATCTACAACTACACCAGCGAGAGAACCAGAGGCACCGCCATC
GACAGCAACGACGTGGACTGCGAGCCCGTGATCACCTACAGCAACATCGGCGTGTGCAAGAACGGCGCCCTGGTG
TTCATCAACGTGACCCACAGCGACAAGACCCACACCTGCCCCCCCTGCCCCGCCCCCGAGCTGCTGGGCGGCCCC
AGCGTGTTCCTGTTCCCCCCCAAGCCCAAGGACACCCTGATGATCAGCAGAACCCCCGAGGTGACCTGCGTGGTG
GTGGACGTGAGCCACGAGGACCCCGAGGTGAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCACAACGCCAAG
ACCAAGCCCAGAGAGGAGCAGTACAACAGCACCTACAGAGTGGTCAGCGTGCTGACCGTGCTGCACCAGGACTGG
CTGAACGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAGGCCCTGCCCGCCCCCATCGAGAAGACCATCAGCAAG
GCCAAGGGCCAGCCCAGAGAGCCCCAGGTGTACACCCTGCCCCCCAGCAGAGAGGAGATGACCAAGAACCAGGTC
TCCCTGACCTGCCTGGTGAAGGGCTTCTACCCCAGCGACATCGCCGTGGAGTGGGAGAGCAACGGCCAGCCCGAG
AACAACTACAAGACCACCCCCCCCGTGCTGGACAGCGACGGCAGCTTCTTCCTGTACAGCAAGCTGACCGTGGAC
AAGAGCAGATGGCAGCAGGGCAACGTGTTCAGCTGCAGCGTGATGCACGAGGCCCTGCACAACCACTACACCCAG
AAGAGCCTGAGCCTGAGCCCCGGCAAGCTCGAGTTTAAAC

Infectious clone cDNA fragments

NNNN Intra-unit overlaps
NNNN Fragment A/C overlaps (incorporating BsmBI site)
NNNN XmaI site
NNNN SalI site
NNNN BsmBI site

Fragment 1A

ACTTGGGCCCTAATACGACTCACTATAG
ACTTTTAAAGTTAAGTGAGTGTAGCGTGGCTATAACTCTTCTTTTACTTTAA
CTAGCCTTGTGCTAGATTTGTCTTCGGACACCAACTCGAACTAAACGAAATATTTGTCTCTCTATGAAAC
CATAGAAGACAAGCGTTGATTATTTCACCAGTTTGGCAATCACTCCTAGGAACGGGGTTGAGAGAACGG
CGCACC
AGGGTTCCGTCCCTGTTTGGTAAGTCGTCTAGTATTAGCTGCGGCGGTTCCGCCCGTCGTAGTTGGGTAG
ACCGGGTTCCGTCCTGTGATCTCCCTCGCCGGCCGCCAGGAGAATGAGTTCCAAACAATTTAAGATCCTC
GTTAATGAGGACTACCAAGTCAACGTGCCTAGTCTTCCTTTCCGTAACGCTTTGCAGGAAATTAAGTACT
GCTACCGTAACGGTTTTGAGGGCTATGTTTTCGTACCTGAATACCGTCGTGACCTAGTTGATTGCGATCG
TAAGGATCACTACGTCATTGGTGTTCTGGGTAACGGAATAAGTGATCTTAAACCTGTTCTCCTTACCGAA
CCTTCCGTCATGTTGCAAGGCTTTATTGTTAGAGCTGGCTGCAATGGCGTTCTTGAGGACTTTGACCTTA
AAATCGCCCGTACTGGAAATGGCGCCATATATGTAGATCAATACATGTGTGGTGCGGATGGAAAACCAGT
TATTGAAGGTGAGTTTAAGGATTATTTCGGTGATGAGGACGTTATCATCTATGAAGGAGAGGAGTACCAT
TGTGCCTGGCTAACAGTGCGCGATGAAAAACCTTTATGTCAACAAAATCTCCTCACAATTAGGGAAATCC
AATACAATCTGGACATCCCACATAAGTTGCCTAATTGTGCTATCAGAGAGGTAGCACCACCAGTCAAGAA
GAATTCCAAGGTTGTTCTTTCTGAAGAGTATAAAAAGCTCTATGACATCTTTGGTTCACCATTTATGGGC
AATGGGGACAGTCTTAACAAGTGTTTTGACAGCCTCCACTTCATTGCTGCCACTCTTAAGTGTCCTTGTG
GCGCCGAGAGTAGTGGTGTTGGTGATTGGACTGGTTTTAAGACTGCTTGTTGTGGTCTTCAAGGCAAAGT
TAAAGGCGTCACATTAGGTGCTGTTAAACCAGGTGATGCCGTTGTCACTAGCATGAGTGCAGGTAAGGGT
GTAAAATTCTTTGCTAATAGTGTACTCCAATATGCTGGTGATGTGGAAAATGTCTCTGTTTGGAAAGTGA
TTAAAACTTTCACTGTTAATGAGACTGTCTGCACCGCTGATTTTGAAGGCGAATTGAACGACTTCATTAA
ACCAGAGAGTACTTCACCTGTTTCATGTTCTATTAAGAGAGCGTTCATCACTGGTGAGGTTGATGATGCA
GTCCATGACTGTATTATCTCAGGAAAGTTGGACCTTAGTACTAACCTTTTTGGTGGTGCTAACCTGCTTT
TCAAGAAAACGCCCTGGTTTGTGTTAAAGTGTGGTGCTATTTTTGCTGATGCTTGGAAAGTTGTTGAAGA
ACTGCTTTGTTCTCTCAAACTTACATACAAGCAAATCTATGATGTAGTAGCATCACTTTGCACTTCTGCC
TTTACTATTATGGATTACAAACCTGTATTTGTAGTTGCATCTAATAGTGTTAAGGATCTCGTTGACAAGT
GTGTCAAAATTCCCCGGGTTTAGCATTACAAAGGCTTGAGACG

Fragment 1B

GTCGACCTCGTTGACAAGTGTGTCAAAATTC
TTGTGAAAGCCTTTGATGTTTTCACACAGACAATTACCATTGCGGGTGTGGAGGCTAA
GTGCTTTGTGCTTGGCTCTAAATACTTGCTTTTTAATAATGCACTTGTTAAGCTTGTCAGCGTTAAAATA
CTTGGCAAGAGGCAGAAAGGTCTTGATAGTGCATTCTTTGCCACTAATTTGATTGGTGCTACTGTTAAGG
TGACACCACAAAGAACTGAGGCAGCCAATATTAGTTTGAATAAAGTAGATGATGTTGTCACACCTGGTGA
AGGTCATATTGTCATTGTAGGCGACATGGCTTTTTACAAGAGTGGAGAGTACTACTTTATGATGGCAAGT
CCTGATTCAGTTCTCATTAATAACGTTTTTAAAGCAGCTCGAGTGCCATCTTATAACATTGTTTACAATG
TGGATGATGATACCAAGAGTAAGATGTCATCTAAAATTGGAACGCCATTTGAATTTGATGGTGATGTGGA
TGCTGCTATTGTTAAGATTAACGACTTACTTGTTGAGTTTAGACAAGAAAAGTTTTGTTTCAGAGCACTT
AAGGATGGTGAAAGTATCTTGGTTGAGGCATACCTTAAGAAATATAAAATGCCAATTTGTCTCAAAAATC
ATGTTGGTTTATGGGACATCATAAGACGCGATTCTGATAAGAAGGGTTTTCTTGACACTTTCAGTCATCT
TAATGAATTGGAGGATGTCAAGGATGTAAAAATTCAGTCTATTAAAAATATCCTTTGCCCTGACCTTCTG
TTAGAATTAGATTTTGGTGCAATTTGGTATAGATGTATGCCAGCCTGCTCTGATGTGTCTATTTTGGGTA
ATGTTAAAATCATGCTAGGTAATGGTGTTAAAGTGATTTGTGATGGCTGCCATGGTTTTGCTAACCGTTT
AACGGTTAATTACAGCAAACTGTGTGACACAGCTCGTAAAGACATTGAAATTGGTGGTATACCATTTTCC

ACCTTTAAAACACCATCTAGCAGCTTCATTGACATGAAAGATGCCATCTATTCTGTGGTTGAATACGGTG
AAGCATTTTCCTTTAAGACTGCAAGTGTACCTGTAACTAATAGTGGTACTGTCAAAAATGACGACTGGTC
AGACCCCATTTTGTTAGAACCTGCAGATTATGTCGAACCTAAAGATAACGGTGATGTAATTGTCATTGCT
GGATACACCTTTTACAGGGATGAGAACGATCACTTTTATCCATATGGTTCCGGCATGGTTGTGCAGAAGA
TGTACAACAAGATGGGTGGTGGTGACAAGTCAGTTTCATTTTCAGACAACGTTAATGTTAAAGAGATAGA
ACCTATTACACGTGTCATGCTAGAATTTGAATTTGACAATGAGGTGGTAACACAGGTTCTTGAAAAAGTC
ATTGGTGCCAAATATAAGTTTACAGGTACCACTTGGGAAGAATTTGAAGATTCTATTTCTGAGAAACTTG
ACAAGGTTTTCGACACGCTTGCTGAACAGGGTGTAGAACTTGAAGGTTACTTCATTTATGACACTTGTGG
CGGATTTGACATTAATAATCCAGATGGTGTTATGATTTCACAGTATGATCTTAACTCAGCTGTTGATGTC
AAAAGTGATTCTGATGCAAGCGTGGAAGACACATCTTCGATTTCTGATAACGAAGATGTTGAACAAATAG
AAAAAGAGAACGTCTCAGCTGCTGATGTTACAGTTGTAGATGTTGAAGAATTTGTGTCGAC

Fragment 1C

GTCGACCAGTTGTAGATGTTGAAGAATTTGTTGAGCAAGTTTCTCT
TGTTGACACCGAAGAACAACTCCCCTCAGTAGAGGAAAAAGTTGAAGGCTCTGCTAAAGATGACCCGTGG
GCTGCTGCTGTTGATGAACAGGAAGCTGAACAACCTAAACCTTCCTTAACTCCATTTAAGACAACAAATC
TTAATGGTAGAATTATTCTTAAACAACAAGATAACAATTGTTGGATAAACGCTTGCTGTTACCAGTTGCA
AGCTTTTGACTTTTTCAACCATGAATTATGGGATGGTTTTAAGAAAGATGACGTTATGCCTTTCGTAAAC
TTCTGTTATGCAGCTCTAACTTTGAAGCAGGGTGATTCAGGAGACTCAGAATACCTTTTGGAGACGATGC
TCAATGATTATAGTACAGCTAAAGTTGTTCTTAGTGCTAAGTGTGGCTGTGGTGTTAAAGAAATAGTTTT
AGAGCGAACTGTGTTTAAGCTCACACCTCTTAAGAATGAGTTTAAATATGGTGTCTGTGGTGACTGTAAA
CAAATTAACATGTGTAAGTTTGCAAGCGTGGAAGGTTCAGGTGTTTTTGTTCATGATAGAATTGAAAAAC
GAACACCTGTGTCACACTTCATTGTGACACCTATTATGCACGCAGTTTACACAGGTACAACACAGAGTGG
TCATTACATGATTGAGGATTGCACTCATGATTATTGTGTGGATGGTATGGGAGTTAAACCCCGTAAACAA
AAGTTTTACACATCCACCTTGTTTCTTAATGCAAATGTAATGACAGCTGACTCTAAAACTACAATCGAAC
CAACTGCACCTGTTAAGGATAAGCTTGTAGAAGAATGTCAAAGTCCTAAAGACTTTATATTACCCTTTTA
CACAGCGGGGAAAGTGTCATTTTACCAAGGAGATTTGGATGTTTTGATTAACCTTCTTGAACCTGACGTT
GTTGTTAACGCTGCTAATGGTGATCTTAGGCACATGGGTGGTGTTGCTAAAGCCATTGATGTGTTCACAG
GTGGTAAGTTGACGAAACGTTCTAAAGAATATCTTAGGACTAATAAAACAATAGCTCCTGGTAATGCTGT
GCTTTTTGAAAACGTACTTGAGCATCTTAGTGTATTGAATGCTGTCGGACCACGTAATGGTGACAGTAGA
GTTGAAGGTAAACTCTGTAACGTTTACAAAGCAATTGCCAAGTGTGACGGTAAGATATTAACACCACTTA
TTAGTGTTGGTATCTTCAAAGTCAAACTGGAAGTTTCATTGCAGTGTCTACTTAAAACTGTCACAGATAG
GGAATTAAATGTCTTTGTTTACACTGACCAAGAGAGAATTACCATAGAAAATTTCTTCAATGGTACTATT
CCTGTCAAGGTAACGGAAGACACTGTGAATCAAAAGCGTGTGTCTGTGGCACTTGATAAAACTTATGGTG
AGCAACTCAAAGGCACTGTTGTTATTAAAGACAAAGATGTTACTAACCAGTTGCCAAGTGTGTCAGATGT
TGGTGAAAAAGTCGTTAAGGCATTGGATGTTGATTGGAATGCTTACTATGGCTTTTCAAATGCAGCGGCT
TTTAGTGCCAGTTCCCATGATGCTTATATATTTGAAGTTGTAACACACAACAACTTCATTGTTCATAAGC
AGACTGATAATAATTGTTGGGTTAATGCAATTTGTTTAGCATTACAAAGGCTTGAGACGGTCGAC

Fragment 2A

CGTCTCCGGCTTAAACCAACATGGAAGTT
TCCTGGTGTTAAAAGTTTGTGGGATGATTTTCTTGTTCGTAAAACTGCTGGATTTGTCCACATGCTATAC
CACATTTCAGGCTTAAAGAAAGGACAACCTGGTGATGCTGAGTTGACTTTACATAAGCTTGGTGAGTTGA
TGTTGAGTGATAGTGCTGTCACCATCACACATACAACTGCATGTGACAAGTGTGCAAAAACTGAAACTTT
TACAGGCCCTGTGATAGCAGCACCTCTTCTGATCCATGGCACTGATGAAACCTGTGTTCATGGTGTAAGT
GTTAATGTTAAAGTGACAAGTGTTAGAGGCACTGTGGCTATCACTTCCTTAATCGGTCCTGTTGTAGGAG
ATGTTATTGATGCAACTGGTTATATTTGTTATACTGGATTGAACTCACGTGGTCATTATACCTATTATGA
TAATCTTAATAGTTTAATGGTTGATGCTGAAAAGGCCTATCATTTTGAGAAAAACCTCTTACAGGTTACA
ACAGCCATTGCTAGTAATTTTATTGCCTACACACCTAAGAGTGAGATCTATCCTAAGGTTCAACCAGTTA
AAGAATCACATACAGCTCGTGTTTTTAGTGAGGTTGAAGAAATTCCTAAAAATGTAGTTCGTGAAACAAA
ATTGCTAGCTATTGAGAGTGGTGCAGATTACACACTCTCAACACTTGGTAAATATGCTGATGTTTTCTTT
ATGACCGGTGATAAAATTCTTAGGTTTCTCCTTGAAGTCTTTAAATATTTATTGGTTGTTTTTATGTGTC
TTAGAAAATCTAAGATGCCTAAAGTTAAGGTAAAACCACCGCATGTATTCAAAAACTTAGGTGCTAAAGT
TAGGACATTGAATTACGTAAGACAATTGAACAAACCAGCGTTATGGCGTTACGTAAAACTAGTGTTACTT

TTAATAGCATTGTACCACTTTTTCTACTTGTTTGTTAGTATACCAGTAATGCACAAATTAGTGTGTAGTA
GTAGTGTACAAGCGTATAGTAATTCTAGTTTTGTGAAATCAGAAGTTTGTGGTAATTCTATTCTATGCAA
GGCGTGTCTGGCATCTTATGATGAATTGGCTGATTTTGATCATCTTCAGGTGTCATGGGATTATAAGTCG
GATCCGCTATGGAATAGAGTTATACAATTGTCTTACTTCACATTCTTGGTTGTTTTTGGCAATAATTATG
TTAGATGTCTTCTTATGTATTTTATCTCTCAGTACCTTAACTTGTGGTTGTCTTATTTTGGTTATGTTAA
GTATAGTTGGTTTTTGCATGTCATCAACTTTGAATCAATTTCTGTTGAATTTGTGATCATAGTTGTAGTT
TTTAAGGCAGTTCTCGCACTTAAACACATCTTCTTTTCATGCAATAACCCATCATGTAAGACTTGTTCTA
AGATTGCTAGGCAGACACGTATTCCTATACAAGTTGTGGTTAATGGTTCAATGAAGACTGTGTATGTCCA
CGCTAATGGTACTGGTAAACTCTGCAAAAAGCATAATTTCTACTGTAAAAATTGTGACTCTTACGGTTTT
GACCACACATTCATCTGTGACGAGATTGTGCGTGACCTGGGTAACAGTATCAAGCAAACTGTTTATGCCA
CTGATAGATCATACCAGGAAGTCACAAAGGTTGAATGCTCTGATGGCTTTTACAGGTTTTATGTTG
CCCGGGACCAGTAGATATGCAAGGCGAGACG

Fragment 2B

GTCGACGATGGCTTTTACAGGTTTTATGTTGGTGA
TGAATTCACGGCATACGACTATGATGTAAAAAACAAGAAGTATAGTAGCCAGGAAGTGCTTAAAACTATG
TTCCTGCTTGATGACTTCATCGTCTATAGTCCATCAGGGTCTTCCCTCGCTAGTGTTAGAAATGTCTGTG
TTTATTTTTCACAACTGATTGGTAGGCCTATTAAAATTGTTAATAGTGAACTTCTTGAGGATTTATCTGT
AGACTTTAAGGGAGCACTTTTTAATGCTAAGAAAAATGTTATTAAAAACTCTTTTAATGTTGATGTGTCA
GAATGTAAAAACCTTGAAGAGTGTTATAAAGCATGCAACCTTGATGTGACATTTTCTACTTTTGAAATGG
CTGTTAATAACGCTCATAGATTTGGTATATTGATTACAGACCGCTCCTTTAACAATTTTTGGCCATCAAA
AATTAAGCCAGGTTCTTCTGGTGTCTCTGCTATGGACATTGGTAAGTGCATGACTTTTGATGCAAAGATT
GTTAATGCAAAGGTTTTAACGCAACGAGGTAAAAGTGTTGTATGGCTTAGTCAGGACTTTTCTACTCTTA
GTTCTACAGCACAGAAAGTTTTAGTTAAAACTTTTGTAGAAGAAGGTGTGAACTTTTCACTAACGTTTAA
TGCTGTGGGTTCAGATGAAGACCTCCCCTATGAGAGATTCACTGAATCAGTTTCACCAAAGAAAGGTTCA
GGTTTCTTTGATGTACTTAAACAGCTTAAACAGATCTTCTGGTGTTTGGTTTTGTTTATTACCTTGTATG
GTTTATGCTCAGTCTACAGTGTTGCAACTCAGTCTTATATTGATTCTGCTGATGGTTATGAATACATGGT
TATTAAGAATGGCATTGTTCAACCATTTGATGATTCTATTAACTGTGTACATAACACATACAAAGGCTTT
GTTATGTGGTTTAAAGCTAAATATGGTTTTGTACCTACATTCGACAAGTCTTGTCCTATTGTTTTAGGAA
CTGTTTTTGACTTAGGCAATATGAGACCAATCCCTGATGTGCCAGCTTATGTAGCACTTGTTGGTAGATC
CCTAGTGTTTGCGATTAATGCAGCATTTGGTGTTACTAATGTTTGTTATGATCATACAGGTTCTGCTGTG
AGTAAAAACTCTTATTTTGATACCTGTGTATTTAATGCTGCTTGCACCACTCTATCAGGCCTTGGAGGTA
CAATTGTTTATTGTGCCAAACAAGGACTTGTGGATGGTGCTAAACTTTATAGCGAATTGATGCCTGACTA
TTTTTATGAGCACGCTAGTGGTAACATGGTTAAAATACCAGCCATTATTAGGGGTTTTGGTCTGAGATTC
GTAAAAACACAGGCTACAACTTATTGCAGAGTAGGGGAGTGTGTTGAAAGTCAGGCCGGTTTTTGCTTTG
GTGGTGACAACTGGTTTGTTTATGATAAAGAGTTTGGTGATGGTTATATTTGTGGAAGCTCTACATTGGG
TTTTCTCAAAAATGTCTTTGCTCTCTTTAACTCTAACATGTCTGTAGTAGCCACATCCGGTGCTATGTTG
GTTAACATTATCATCGCATGTTTGGCTATAGCGGTCTGTTATGGTGTTCTTAAATTTAAGAAAATCTTTG
GTGATTGTACTCTCCTAGTTGTTATGATTATTGTTACATTGGTTGTTAACAATGTTTCT
TATTTTGTGACCCAAAACACATTCTTTATGATTGTTTATGCCATTATTTAGTCGAC

Fragment 2C

GTCGACTTATGATTGTTTATGCCATTATTTATTACTTTACAACAAGGAAATTATCATACCCT
GGTATCTTGGATGCAGGATTCATTATTGCCTATGTGAACATGGCCCCATGGTATGTGCTTGTGCTGTACA
TAGTGGTCTTCTTATATGATTCTCTACCCTCACTATTCAAACTTAAGGTTACAACAAACCTTTTTGAAGG
GGATAAATTTGTTGGTAGTTTTGAATCAGCAGCTATGGGTACTTTTGTTATAGATATGCGTTCATATGAG
ACACTTGTTAATTCTACATCCTTGGATAGAATTAAGTCATATGCTAACAGCTTTAACAAGTACAAGTACT
ATACTGGTTCAATGGGAGAGGCCGATTACCGCATGGCTTGTTATGCTCATTTAGGTAAGGCATTAATGGA
TTATTCAGTCTCAAGAAATGATGTTCTTTACACACCACCAACGGTCAGTGTTAATTCAACATTACAGTCA
GGACTGAGAAAAATGGCACAGCCTAGTGGTATTGTGGAACCCTGCATTGTGAGGGTAGCCTATGGCAATA
ATGTGCTTAATGGTTTGTGGCTTGGAGATGAAGTTATCTGCCCTAGACACGTCATTGCAAGTGACACATC
ACGTGTGATCAATTATGACAATGAGTTGTCCGGTGTGCGGTTACATAACTTCTCTATAGCCAAAAATAAT
GTGTTTTTGGGCGTTGTGTCTGCCAAGTATAAGGGTGTGAACCTTGTGCTTAAGGTTAATCAGGTAAACC
CCAACACACCAGAACACAAATTCAAATCCGTGAAGTCTGGTGAGAGTTTTAACATTCTCGCTTGTTATGA

AGGCTGTCCCGGTAGTGTCTACGGTGTGAACATGAGAAGTCAGGGTACCATCAAAGGTTCATTTATTGCT
GGTACCTGTGGGTCAGTAGGTTATGTATTAGAAAATGGAATACTCTATTTTGTGTACATGCACCACTTAG
AATTAGGTAATGGCTCTCATGTTGGTTCAAACCTTGAGGGGGTAATGTATGGCGGTTATGAAGATCAACC
TAGCATGCAATTGGAGGGTACTAACGTCATGTCATCAGATAATGTAGTTGCATTTTTGTATGCTGCTCTT
ATTAATGGTGAAAGATGGTTTGTTACAAATGCATCAATGACGTTAGAATCTTACAATGTATGGGCCAAAA
CTAACAGTTTTACGGAACTTGTTTCAACTGATGCTTTTAATATGTTGGCTGCAAAAACTGGTTATAGTGT
TGAGAAGTTGCTTGAATGTATTGTTAGACTCAACAAAGGTTTTGGAGGACGCACTATACTCTCTTATGGT
TCATTGTGTGATGAATTCACACCTACGGAAGTTATAAGGCAAATGTATGGTGTTAATCTTCAGAGTGGTA
AAGTGAAATCTCTCTTTTACCCTGTGTTGACAGCCATGGCTATTCTATTTGCTTTTTGGTTAGAGTTCTT
TATGTACACACCCTTTACATGGATTAATCCATCGTTTATCAGTGTTATTCTAGCTATTACAACTTTTTTC
TCTGTGTTATTAGTTGTTGGGATTAAACACAAAATGTTATTCTTTATGTCTTTTGTAATGCCTAGTGTTG
TGTTGGCCACCGTTCATAATGTTGTATGGGATATGACTTATTATGAAAGCTTACAAGCTCTTGTGGCGAA
CACTAACACAACATTTTTACCAGTAGATATGCAAGGCGAGACGGTCGAC

Fragment 3A

CGTCTCCAAGGCGTTATGCTTGCTTTGTTCTGTGTTGTTGTGTTT
GTTACATACACTATAAGATTCTTTACATGCAAGCAATCCTGGTTTTCTTTATTTGTAACAACTGTCTTTG
TGGTGTTCAACATTGTTAAGTTGCTAGGCATGGTAGGTGAACCATGGACTGAAGATCACATCTTGTTATG
TCTTGTAAACATGCTTACCATGTTGATAAGTCTTACTACAAAGGACTGGTTTGTTGTTTTTGTATCATAC
AAATTTGCTTATTATGTCGTTGTTTATGTAATGCAACCATCCTTTGTTCAAGACTTTGGGTTTATTAAGT
GTGTAAGCGTGATTTACATGGCATGTGGTTATTTGTTTTGCTGTTATTATGGTATTCTTTACTGGGTTAA
CAGATTTACATGCATGACGTGTGGTGTTTACCAATTCACTGTCTCACCGGCAGAGCTCAAGTATATGACT
GCTAACAATTTGTCAGCACCTAAGACCGCATATGATGCCATGATTCTTAGTGTTAAATTGATGGGTATAG
GTGGAGAGAGAAATATAAAAATCTCCACTGTACAGTCTAAACTCACAGAGATGAAATGCACAAACGTTGT
GCTGCTTGGGCTTTTGTCTAAAATGCATGTTGAATCTAACTCAAAGGAGTGGAATTATTGTGTTGGGCTG
CATAATGAGATTAATTTGTGTGATGACCCCGAAGTGGTTCTTGAAAAACTACTTGCACTCATCGCATTCT
TTTTATCCAAACACAATACATGTGATCTTAGTGACCTTATCGACTCCTATTTTGATAACACCACTATTTT
ACAGAGTGTGGCTTCTGCTTATGCTGCTTTACCTAGCTGGATTGCATATGAAAAAGCTCGCGCGGATTTA
GAAGAGGCTAAGAAAAATGATGTTAGTCCACAATTATTGAAACAGCTTACAAAAGCATGCAATATTGCTA
AGAGTGAGTTTGAACGCGAGTCCTCGGTTCAAAAGAAGCTTGATAAAATGGCCGAGCAAGCTGCAGCTAG
TATGTACAAAGAGGCAAGGGCTGTTGACAGAAAATCAAAAATAGTTTCTGCTATGCATAGTCTTCTTTTT
GGTATGCTCAAGAAACTTGATATGTCCAGTGTTAATACTATTATTGAACAGGCCCGTAATGGTGTGTTGC
CCCTGAGTATCATACCTGCTGCAGCAGCCACAAGACTTATTGTTGTAACACCTAATATTGAAGTGCTTTC
CAAAATTAGGCAAGAAAACAATGTCCACTATGCAGGTGCTATCTGGTCAATTATTGAAGTCAAGGATGCT
AATGGTGCTCAAGTCCATTTAAAAGAAGTCACAGCTGCCAATGAGTTGAACCTGACCTGGCCTTTGAGTA
TTACATGTGAGAGAACCACAAAGTTGCAAAACAATGAAATTATGCCAGGTAAACTTAAAGAGAAGGCTGT
TAAAGCGTCTGCTACAATTGATGGTGATGCTTATGGTAGTGGTAAGGCCCTCATGGCTTCTGAAAGTGGT
AAGAGCTTTATATATGCATTTATAGCGTCTGACAGCAACCTTAAGTATGTTAAATGGGAGAGCAATAATG
ATGTAATACCCATTGAACTTGAGGCCCCATTGCGCTTTTATGTAGATGGTGCTAATGGCCCTGAAGTTAA
GTATCTTTACTTTGTGAAGAACTTGAACACTCTTAGACGTGGTGCTGTTC
CCCGGGGGTTGTACTGTTAATGATGGAGACG

Fragment 3B

GTCGACAACACTCTTAGACGTGGTGCTGTTCTGGGTTATATTGGCGCTACT
GTTCGTTTGCAGGCTGGTAAACCTACTGAACACCCATCTAACAGTGGTTTATTAACCTTGTGTGCTTTCG
CTCCTGACCCTGCTAAGGCATACGTTGATGCTGTTAAGAGAGGTATGCAGCCAGTTACTAATTGTGTGAA
AATGCTCTCAAATGGTGCTGGTAATGGTATGGCTATTACAAATGGTGTAGAGTCTAATACGCACCAAGAC
TCTTATGGTGGTGCATCTGTGTGCATCTATTGTAGGTGTCATGTAGAACATCCTGCAATTGATGGATTGT
GTCGTTATAAGGGTAAATTTGTTCAAGTACCAACCGGCACTCAAGATCCTATTAGATTTTGTATTGAAAA
TGAAGTCTGTGTTGTTTGTGGTTGCTGGCTTAATAATGGTTGCATGTGTGATCGTACTTCTATGCAAGGC
ACAACTGTTGACCAGAGTTATTTAAACGAGTGCGGGGTTCTAGTGCAGCTCGACTAGAACCCTGTAATGG
TACTGATCCAGACCACGTTAGTAGAGCTTTTGACATCTATAATAAAGATGTTGCATGTATTGGTAAATTC
CTTAAGACTAATTGTTCCCGTTTTAGGAATTTGGACAAACGTGATGCCTACTATGTTGTCAAGCGTTGTA
CTAAGAGCGTTATGGACCATGAGCAAGTCTGTTATAACGATCTTAAAGATTCTGGTGCAGTTGCTGAACA

TGACTTCTTTTTGTATAAGGAAGGTCGATGTGAATTCGGCAACGTCGCACGTAAGGATCTTACAAAGTAC
ACTATGATGGATCTATGTTATGCCATTAGGAATTTTGATGAAAAGAATTGCGAAGTTCTTAAAGAAATAC
TCGTGACATTAGGTGCCTGCAATGAATCCTTCTTTGAAAATAAAGATTGGTTTGACCCAGTGGAAAATGA
AGCTATACATGAGGTGTACGCAAGACTTGGACCTATAGTAGCCAACGCTATGCTTAAATGTGTAGCGTTT
TGTGATGCTATCGTGGAGAAAGGTTACATAGGTATTATTACACTTGATAACCAAGACCTTAATGGTAATT
TTTACGACTTCGGTGATTTCGTAAAGACAGCGCCTGGTTTTGGATGTGCTTGCGTGACGTCATATTATTC
TTATATGATGCCCTTAATGGGAATGACTTCATGTTTAGAGTCTGAAAATTTTGTGAAAAGTGACATTTAT
GGTTCTGATTATAAGCAGTATGATTTGCTAGCTTATGATTTTACAGACCATAAAGAAAAGCTGTTCGAAA
AATATTTCAAACACTGGGACCGCACATACCATCCTAACTGCTCTGATTGTACAAGTGATGATTGTATCAT
TCATTGTGCTAATTTTAACACATTGTTTTCAATGACCATACCCAACACTGCTTTTGGACCTCTTGTACGT
AAAGTCCACATAGATGGTGTGCCAGTTGTTGTTACTGCTGGTTACCATTTTAAACAGTTGGGTATTGTGT
GGAATCTTGATGTTAAATTAGACACTATGAAGTTGACTATGACAGATCTTCTTAGGTTCGTCACTGACCC
TACATTACTTGTAGCATCAAGCCCTGCATTGTTAGATCAGCGCACTGTCTGTTTTTCCATTGCAGCTTTG
AGTACTGGTGTTACATATCAGACAGTGAAACCAGGTCACTTTAACAAGGATTTCTACGATTTTATAACAG
AGCGTGGATTCTTTGAGGAGGGAGTCGAC

Fragment 3C

GTCGACAG
AGCGTGGATTCTTTGAGGAGGGATCTGAGTTGACGTTGAAACATTTCTTCTTTGCACAGGGTGGTGAAGC
CGCTATGACAGATTTTAACTACTACCGTTATAACAGAGTTACTGTTTTAGACATCTGTCAGGCGCAATTT
GTTTATAAAATTGTGTGTAAGTACTTTGATTGTTACGATGGTGGCTGCATCAATGCTCGAGAGGTTGTTG
TTACCAACTACGACAAGAGTGCTGGTTATCCACTTAATAAATTTGGTAAGGCAAGACTTTACTATGAGAC
TCTGTCATATGAGGAGCAGGATGCGATTTTTGCATTAACAAAGAGAAATGTTCTACCCACAATGACACAA
ATGAATTTGAAATATGCTATATCTGGAAAGGCTAGAGCTCGCACTGTAGGAGGAGTGTCACTCCTTTCGA
CAATGACTACGAGACAATACCACCAGAAGCATTTAAAGTCTATTGCTGCTACACGTAATGCCACTGTTGT
TATTGGAACGACTAAGTTCTACGGTGGCTGGGATAATATGTTAAAGAATTTGATGCGTGATGTAGATAAC
GGCTGTTTGATGGGATGGGACTATCCAAAGTGCGACCGTGCTTTGCCTAATATGATACGAATGGCTTCTG
CAATGGTATTGGGTTCTAAGCACATTGGGTGTTGTACACATAGTGATAGGTTCTATCGCCTCTCTAACGA
GTTAGCTCAAGTACTTACAGAGGTTGTGCATTGCACAGGTGGTTTTTATATCAAACCTGGTGGTACAACT
AGTGGTGATGGTACTACAGCATATGCTAATTCGGCTTTTAACATCTTTCAAGCTGTTTCTGCTAATGTTA
ATAAGCTTTTGGGAGTTGACTCAAACACCTGTAACAATGTTACAGTTAAGTCCATACAACGCAAGATATA
TGATAATTGTTATCGAAGTAGCAGTGTTGATGATGACTTCGTTGTTGAGTATTTTAGCTATTTAAGAAAG
CACTTTTCTATGATGATTTTGTCAGATGATGGCGTGGTTTGTTACAACAAAGACTATGCTGATCTTGGTT
ATGTAGCAGACATTAGTGCTTTTAAAGCTACTTTGTATTACCAGAATAATGTTTTTATGTCTACAGCCAA
GTGTTGGGTAGAACCAGACCTTAATGTTGGACCACACGAATTTTGTTCACAACACACACTACAAATTGTA
GGACCGGATGGTGATTATTATTTACCATACCCAGACCCTTCTAGAATCTTGTCAGCAGGTGTTTTCGTTG
ATGACATCGTTAAAACAGATAACGTCATTATGCTTGAACGTTATGTGTCTTTAGCAATCGATGCGTATCC
ACTTACAAAACACCCTAAACCCGCATATCAAAAAGTGTTTTATGCATTGTTAGATTGGGTTAAACATTTA
CAGAAGACTCTAAATGCTGGTATACTTGATTCATTTTCTGTCACTATGTTGGAGGATGGTCAAGATAAAT
TCTGGAGTGAAGAGTTCTATGCCAGTCTTTACGAAAAGTCTACTGTTTTGCAAGCTGCCGGTATGTGTAT
TGTTTGTGGTTCACAAACTGTGTTACGTTGCGGAGACTGTTTAAGGAGACCTCTTTTATGCACTAAGTGC
GCTTACGACCATGTCATGGGTACAAAGCATAAATTCATTATGTCTATCACACCATATGTGTGTAGCTATA
ATGGTTGTACTGTTAATGATGGAGACGGTCGAC

Fragment 4A

CGTCTCCTGATGTAACAAAATTGTTTTTGGGAGGTCTTAACTATTATTGTACTGAACATAA
ACCACAATTGTCATTCCCGCTCTGTGCTAATGGCAATGTGTTTGGTCTGTACAAGAGTAGTGCACTTGGT
TCTGAGGACGTTGACGATTTCAATAAACTTGCAATCTCGGACTGGACCAATGTAGAAGATTACAAACTTG
CTAATAATGTCAAAGAATCTCTGAAAATTTTCGCTGCTGAAACTGTGAAAGCAAAGGAGGAGTCTGTTAA
AGCTGAATATGCTTATGCCATATTAAAGGAGGTGGTCGGCCCTAAGGAAGTTGTACTCCAATGGGAATCC
TCTAAGGTTAAACCTCCACTTAACAGAAATTCTGTTTTCACATGTTTTCAGCTAAACAAGGATACTAAAA
TTCAGTTAGGTGAATTTGTGTTTGAGCAGTCAGAGTATGGTAGTGAATCTGTTTACTACAAGAGTACTAG
CACTACTAAGCTGACACCGGGTATGATTTTTGTGTTGACATCTCATAACGTGAGTCCACTTAAAGCTAAT
ATCTTAGTCAACCAAGAGAAGTACAATACCATATCCAAGCTCTATCCTACGTTCAACATAGCGGAAGCCT

ATACCACATTGGTGCCTTACTATCAAATGATTGGTAAGCAAAAGTTTACAACTATCCAAGGTCCTCCTGG
TAGTGGTAAATCACATTGTGTTATAGGTTTGGGTTTGTATTATCCTCAAGCTAGGATCGTCTACACAGCA
TGTTCACATGCTGCTGTGGATGCATTGTGTGAAAAGGCTTCTAAGAACTTTAATGTGGATAAGTGTTCAA
GGATAATACCCCAAAGAATCAGAGTTGACTGCTATATGGGTTTTAAACCTAATAATACTAATGCACAGTA
CTTGTTTTGTACAGTTAATGCTTTACCTGAAGCTAATTGCGACATCGTAGTTGTGGATGAAGTCTCAATG
TGTACAAATTATGATCTCAGTGTTATAAATAGTAGACTGAGTTACAAGCACATTGTTTACGTGGGAGACC
CACAACAGCTTCCAGCTCCTAGAACCTTGATTAATAAGGGTACACTCCAGCCTGAAGATTACAATGTTGT
AACCCAGAGAATGTGTAAACTAGGACCTGATGTATTTTTGCATAAATGCTATAGATGCCCAGCTGAAATT
GTTAAGACAGTCTCCGCGCTTGTTTATGAGAATAAGTTTTTGCCTGTCAACCCAGAGTCAAAGCAGTGCT
TTAAAATGTTTGTAAAAGGTCAGGTGCAGATTGAATCCAACTCTTCTATAAACAATAAGCAACTAGATGT
TGTTAAGGCATTCCTAGCACATAATCCAAAATGGCGTAAAGCCGTTTTTATCTCACCCTATAACAGTCAG
AACTATGTCGCAAAACGTCTACTCGGTTTGCAAACACAAACGGTAGACTCTGCTCAAGGTAGTGAGTATG
ATTATGTCATTTACACACAGACTTCTGATACGCAGCATGCTATTAATGTCAACAGGTTTAATGTTGCCAT
TACTAGAGCAAAGGTTGGTATTCTCTGTATCATGTGTGATAGAAAGATGTATGACAATCTTGATTTCTAT
GAACTCAAAGATTCAAAAGTTGGCTTGCAAGCAAAACCTGAAACGTGTGGTTTGTTCAAAGATTGTTCAA
AGACTGATCAATACATACCACCAGCATATGCTACAACATATATGAGTCTGTCTGAAAAT
CCCGGGCTTTGTTAACACACCATGAGAGACG

Fragment 4B

GTCGACAACATATATGAGTCTGTCTGAAAATTTTAAGACAAG
TGATGGTTTAGCTGTTAACATCGGTACAAAGGATGTTAAATATGCTAATGTTATCTCATATATGGGATTC
AGGTTTGAGGCCAATGTACCAGGCTATCACACATTGTTTTGCACACGAGACTTTGCTATGCGTAATGTGA
GATCATGGTTAGGGTTTGATGTCGAAGGTGCACATGTCTGTGGTGATAACATCGGAACTAATGTACCACT
ACAGCTGGGTTTTTCAAATGGTGTAGATTTTGTAGTACAAACTGAGGGATGTGTTGTTACTGAGAAAGGT
AATAGTATTGAGGTTGTAAAGGCAAGAGCACCACCGGGTGAGCAATTTGCACATTTGATTCCACTTATGA
GAAAAGGTCAGTCTTGGCACATTGTTAGACGTCGTATAGTGCAGATGGTTTGTGACTATTTCGACGGTTT
ATCTGACATCCTAATTTTCGTGCTTTGGGCTGGTGGTCTTGAGCTTACAACTATGAGATACTTTGTTAAA
ATTGGAAAACCACAGAAATGTGATTGCGGCAAAATTGCAACGTGTTATAACAGCTCGCAGTGTGTCTATG
CTTGTTTTAAACATGCATTAGGATGTGATTACTTGTACAATCCTTACTGCATTGATATTCAACAATGGGG
TTATACAGGATCCTTGAGCATGAACCATCATGAAGTTTGTAACATTCATAGAAATGAGCATGTAGCTAGT
GGTGATGCTATCATGACTAGGTGCCTTGCTATATATGATTGTTTTGTCAAGCGTGTAGATTGGTCCATTG
TGTACCCATTTATTGATAATGAAGAAAAGATCAATAAAGCTGGTCGCATTGTACAATCACATGTCATGAG
AACTGCTCTTAAGGTTTTCAACCCTGCTGCAATTCACGATGTTGGTAATCCTAAAGGTATTCGTTGTGCT
ACAACATCTATACCATGGTTCTGTTATGACCGTGATCCTATTAATAACAATGTTAGATGTCTGGAGTATG
ATTACATGGTACATGGACAAATGAATGGTTTAATGTTATTTTGGAATTGCAACGTAGACATGTACCCGGA
GTTCTCGATTGTTTGTAGATTTGATACTCGAACTCGCTCAAAATTGTCATTAGAAGGTTGTAATGGTGGT
GCATTGTATGTTAATAACCATGCATTTCACACACCAGCTTATGATAGAAGAGCATTTGCTAAGCTTAAAC
CTATGCCATTCTTTTACTACGATGAGAGTGATTGTGAGTTTATTGATGGACAACCTAATTATGTACCACT
TAAAGCTAATGTCTGCATAACTAAATGCAACATAGGTGGTGCAGTCTGTAAGAAACATGCAGCACTTTAC
AGAGCATATGTTGAAGACTACAACACTTTCATACAAGCTGGTTTTACAATTTGGTGCCCTCAAAATTTTG
ATACATATATGTTGTGGCAAGGTTTTGTTAATAGCAAAGCACTCCAGAGTTTGGAAAACGTAGCCTTCAA
TGTCGTCAAGAAAGGTGCCTACGCTGATTTAAAAGGTGACTTACCAACAGCAGTTATAGCTGATAAGATC
ATGGTGAGGGATGGACCTACTGACAAGTGTATTTTCACAAATAAAACTAGTTTGCCTACGAATGTGGCTT
TTGAGCTGTATGCAAAACGCAAACTTGGACTCACACCTCCGTTAACAATACT
TAGAAATCTAGGTGTTGTCGCAACAGTCGAC

Fragment 4C

GTCGACTAGAAATCTAGGTGTTGTCGCAACA
CATAAATTCGTGTTGTGGGATTATGAAGCCGAATGTCCTTTCTCAAACTTTACTAAGCAAGTG
TGTGCGTACACTGATCTTGATGGTGAAGTTGTAACATGTTTTGATAATAGTATTAGTGGTTCTTTTGAAC
GCTTTACTACTACGAAAGATGCAGTGCTTATTTCTAATAACGCTGTGAGGGGACTCAGTGCCATTAAATT
ACAATATGGCTTTTTGAATGATTTACCTGTAAGTACTGTAGGAAACAAACCTGTCACATGGTACATCTAT
GTGCGCAAGAACGGTGAGTACGTCGAACAGATTGACAGTTATTACACACAGGGACGTACTTTTGAAACCT
TTAAACCTCGTAGTACAATGGAAGAAGACTTTTTTAGTATGGATACTACACTCTTCATCCAAAAGTATGG

TCTTGAGGATTATGGTTTTGAACATGTTGTATTCGGAGATGTTTCTAAAACTACCATTGGTGGTATGCAT
CTTCTTATATCACAAGTGCGCCTAGCAAAAATGGGTTTGTTTTCTGTCCAAGAATTTATGAGTAATTCTG
ACAGTACACTGAAAAGTTGTTGTATAACATACGCTGACGATCCGTCTTCTAAGAATGTGTGCACTTATAT
GGACATACTCTTAGACGATTTTGTGACTATCGTTAAGGGCTTAGATCTTAACGTTGTGTCGAAAGTTGTG
GATGTTATTGTAGATTGTAAGGCATGGAGATGGATGTTGTGGTGTGAGAACTCACAAATTAAAACCTTCT
ATCCACAACTCCAGTCTGCTGAGTGGAATCCTGGCTATAGCATGCCTACACTTTACAAAATACAGCGTAT
GTGTCTCGAACGGTGTAATCTCTACAATTATGGTGCACAAGTGAAGTTACCTGACGGCATTACTACTAAT
GTCGTTAAGTATACCCAGTTGTGTCAATATCTCAATACCACTACACTGTGTGTTCCACATAAAATGCGTG
TTCTGCATTTGGGAGCAGCAGGTGCTAATGGTGTTGCCCCTGGTACCACTGTATTAAGAAGATGGTTGCC
AGACGATGCCATATTGGTTGATAATGATATGAGGGATTACGTTTCCGACGCAGACTTCAGTGTTACAGGT
GACTGTACTAATCTTTATATAGAGGACAAATTTGATTTACTTATCTCTGATTTATATGATGGCTCAACCA
AGTCTATAGACGGTGAAAACATGTCGAAAGATGGCTTCTTTACATATATTAATGGTTTTATTAATGAGAA
GCTAGCGCTTGGAGGTTCTGTAGCCATCAAAATCACTGAATTTAGCTGGAATAAAAGTTTATATGAATTA
ATTCAAAGATTTGAGTATTGGACTGTGTTTTGTACAAGTGTTAATACCTCGTCGTCAGAAGGTTTTCTGA
TAGGTATCAACTACTTAGGACCATACTGTGACAAGGCTATAGTGGATGGAAATATAATGCATGCCAATTA
TATATTTTGGAGAAATTCCACAATTATGGCGTTGTCACACAACTCAGTCCTAGATACTCCAAAATTCAAG
TGTCGTTGCAATAACGCACTTGTTGTTAATTTAAAAGAAAGAGAATTAAATGACATGGTTATTGGATTAC
TAAGGAAGGGTAAAATACTCATTAGAAATAATGGTAAGTTACTAAACTTTGGTAATCACTTTGTTAACAC
ACCATGAGAGACGGTCGAC

Fragment 5A

CGTCTCCCATG
ATACTATTGTTGGTACTTTTTAGTGTTGTAGGAGCTCACGATGCACCTCATGGTGTAACACTAC
CACAGTTTAACACCTCATACAACAATGACAAGTTTGAACTTAATTTCTACAATTTCTTGCAAACTTGGGA
TATACCACCAAACACAGAAACTATTCTGGGTGGTTACCTACCATACTGTGGGACAGGTGCTAACTGTGGT
TGGTACAATTTTGTCTACCAACAAAATGTGGGATCAAATGGTAAATATTCTTATATAAACACACAAAATT
TGAACATACCGAATGTTCATGGTGTTTATTTTGACGTACGTGAGCACAATTCAGATGGTGTGTGGGATCT
ACGCGACCGTGTAGGTTTACTCATAGCCATCCACGGCAAATCACAATATAGCCTACTCATGGTTTTGCAA
GACAATGTGGAAGAGAATCAGCCTCACGTAGCTGTTAAAATCTGCCATTGGAAACCTGGCAACATAAGTT
CAATTCATCAATTTAGTGTTAACCTAGGAGATGGTGGTCAATGTGTGTTCAATCAGAGATTCTCATTGGA
CACCATATTGACAACTAATGACTTCTATGGCTTCCAGTGGACTAATAATTATGTCAACATTTATCTAGGT
GGCACCATTACTAAAGTGTGGGTTGAAAATGATTGGAGTGTTGTTGAATCCAGTATTTCCTACCACTGGA
ACCGAATCAACTATGGATATTATATGCAATTTGTCAACCGCACCACTTATTACGTTTATAATAATACTGG
TGGTGCAAATTATACTCACCTGCAGCTCGAAGAGTGCCACTCTGAGTATTGCGCTGGTTATGCTAAGAAT
GTCTTTGTGCCTATTGATGGTAAAATACCAGATGGCTTCTCTTTTAGTAATTGGTTTTTACTATCAGACA
AGTCCACATTGGTACAGGGTCGTGTTCTTAGTAAACAACCAGTTTTCGTACAATGTCTTAGAGCTGTGCC
ATCATGGTCCAACAACAGTGCCGTGGTGCATTTTAGTAATGATGATTTTTGTCCTAATGTTACGGCAGAG
GTGTTAAGGTTTAATCTGAATTTTAGTGATACTGATGTTTATGTAGCTTCAAATAGTGATGATCAGTTGT
ACTTCACATTTGAAGACAATACTACTGCAGGCGTAGCTTGCTATAGTAGCGCTAATATTACAGATTATAA
ACCAAATACTAATGCTAGCTCACAGATCCCTTTTGGTAAAACCACGCATTCATACTTCTGTTTTGCCAAT
TTCTCAAGCTCTCATGTTACACAATTTTTGGGTATATTGCCACCAACTGTGCGAGAGTTTGCGTTTGGCA
AAGACGGTTCCATTTTTGTTAATGGTTATAAATATTTTAGTTTTCCACCTATCAAGAGTGTTAATTTCTC
TATTAGTTCGGTTGAAAATTTTGGTTTTTGGACTATAGCTTACACCAATTACACCGATGTAATGGTGG
CCCGGGAAAGGTTCACATTCATTAAGAGACG

Fragment 5B

GTCGACACCAATTACACCGATGTAATGGTGGAT
GTTAATGGCACTGGTATTACTAGGTTGTTCTATTGTGACTCACCCATTAATAGAATTAAGTGCCAGCAAC
TGAAGTATGAACTACCAGATGGTTTTTATTCAGCTAGCATGCTTGTTAAAAAGGACTTACCTAAAACATT
TGTAACTATGCCACAGTTTTATAATTGGATGAATGTCACTTTACATGTTGTGTTGAATGATACTGAGAAA
AAGGCTGATATTATTCTGGCTAAGGCTTCTGAATTAGCCTCACTTGCAGACGTTCACTTTGAAATTGCAC
AAGCTAATGGCAGTGTGGTTAACGCCACCAGTCTGTGTGTACAGACAAGACAGTTGGCTCTCTTCTATAA
GTATACTAGCTTACAAGGTTTGTATACTTATTCTAATTTGGTAGAACTGCAGAATTATGACTGCCCATTT
TCGCCACATCAATTTAATAATTATTTGCAGTTTGAAACTTTGTGTTTTGATACAAGCCCGGCTGTTGCAG

GTTGCAAGTGGTCTTTAGTTCATGATGTTAAATGGCGCACACAGTTTGCAACAATTACAATCTCTTATAA
AGATGGTGCCAAGATTACAACTATGCCAAAAGCAAAGCTGGGTTTTCAAGATATCTCTAACATTGTAAAA
GATGAATGCACTGATTATAACATTTACGGATTTCAGGGCACTGGTATCATCAGAAACACCACCACTAGGA
TGGTTGCAGGCCTTTACTACACCTCAATTAGTGGTGACTTGCTTGCATTTAAAAATAGCACTACAGGTGA
AATTTTCACCGTAGTGCCATGTGGTTTAACGGCTCAAGCTGCAGTGATTAATGATGAAATAGTTGGAGTT
ATAACAGCTGTTAATCAAACTGATCTATTTGAGTTTGTAAATCATACACAAGCAAGAAGAGCACGCAGGT
CAACAGGTTCACAGACCGTGCAAACTTATACTATGCCACAATTCTATTACATAACAAAGTGGAATAATGA
CAGCTCTACTAATTGTACATCTGTCATTACGTACTCCTCTTTTGCTATTTGTAACACGGGTGAGATTAAA
TATGTTAATGTCACTCATGTGGAAATTGTAGATGATAGTATAGGTGTGATCAAACCTATCTCAACAGGCA
ACATATCCATACCTAAAAATTTCACTGTCGCAGTTCAAGCTGAATACATTCAGATTCAAGTCAAACCCGT
TGTGGTTGATTGCGCTAAGTATGTTTGTAATGGAAATAGTCATTGCCTCACACTACTTGCACAATACACA
TCTGCTTGTCAGACAATTGAAAATGCCCTTAATTTTGGTGCACGTCTTGAATCGCTAATGCTTAATGATA
TGATCACAGTATCAGATCGTAGTTTGGAGCTTGCAACTGTTGAAAAATTTAACACCACGGTTTTAGGTGC
TGAAAAGTTGGGTGGTTTCTATTTTGATGGTTTGAGAGAGTTGTTACCACCTACAATTGGTAAG
AGGTCTGCTATTGAGGATTTACTGTGTCGAC

Fragment 5C

GTCGACAGGTCT
GCTATTGAGGATTTACTGTTTAATAAGGTGGTAACTAGTGGCCTTGGAACTGTTGATGATGATTATAAAA
AGTGCTCTGCTGGTACAGACGTTGCAGACCTTGTGTGTGCTCAATATTATAATGGCATAATGGTACTACC
TGGTGTTGTGGACCAAAATAAGATGTCTATGTACACCGCTTCCTTAATAGGTGGCATGGCCATGGGTTCT
ATCACATCAGCAGTAGCCGTCCCCTTTGCTATGCAAGTGCAGGCCAGGCTTAATTATGTTGCATTACAAA
CGGATGTCTTACAGGAAAACCAGAAAATACTTGCTAATGCCTTTAATAATGCCATCGGCAACATCACGCT
AGCACTGGGAAAAGTTTCTAATGCCATTACTACTATTTCAGATGGATTTAATACAATGGCTTCAGCACTT
ACTAAGATTCAGAGTGTGGTTAATCATCAGGGTGACGCATTGAGTCAACTTACTAGTCAGTTACAGAAGA
ATTTCCAGGCTATTAGTAGTTCTATTGCTGAAATCTACAATAGACTAGAAAAGGTTGAGGCTGATGCTCA
AGTTGACCGCCTCATTACTGGGCGATTGGCAGCACTTAATGCCTATGTGTCTCAAACTTTAACTCAATAC
TCAGAAGTCAAGGCTAGTAGGCAATTAGCAATGGAGAAGGTTAATGAGTGTGTTAAATCACAGTCTAATA
GGTATGGGTTCTGTGGAAATGGAACACACCTATTCTCTCTAGTCAATTCTGCACCTGATGGTTTACTCTT
CTTTCACACAGTTTTGCTTCCTACAGAATGGGAAGTAGTGACGGCATGGTCAGGAATATGTGTTAATGAC
ACGTATGCATATGTGTTGAAAGACTTTGAATATTCTATTTTCAGCTATAACGGTACGTATATGGTGACTC
CTCGTAATATGTTTCAACCTAGAAAACCTCAGATGAGTGATTTCGTGCAAATTACGAGTTGTGAAGTGAC
TTTTCTGAACACTACATATACGACATTTCAGGAGATTGTGATTGATTATATTGATATCAACAAGACTATC
GCTGATATGCTTGAACAATATAATCCCAATTACACAACACCTGAATTAAACTTACAGCTTGAAATCTTTA
ATCAGACAAAGTTAAACCTTACTGCAGAAATAGACCAATTAGAGCAAAGAGCAGACAACCTTACTAACAT
AGCACATGAGCTACAACAGTACATTGACAATCTTAATAAGACCCTTGTTGACCTCGAATGGCTCAACAGA
GTAGAGACTTATGTAAAATGGCCCTGGTATGTGTGGCTACTAATTGGATTAGTAGTGGTCTTCTGCATAC
CATTGTTACTGTTTTGCTGTCTGAGTACTGGTTGTTGTGGATGCTTTGGTTGTCTTGGCAGTTGTTGTCA
TTCTCTTTGCAGCAGACGACAATTTGAAAGTTATGAACCCATTGAAAAGGTTCACATTC
ATTAAGAGACGGTCGAC

Fragment 6A

CGTCTCCATTAACTAAAC
GATTTATGGACACTGTCAAGTCCATTGGCATCTCTGTAGACGCTGTACTTGACGAGTTAGATTCTATTGC
TTTTGCTGTAACACTTAAAGTATTATTTACTTCTGGTAGATTACTTGTGTGTATAGGTTTTGGTGATACT
TTTGAGGAGGCTGAACAAAAAGCTTATGCCAAACTTCAGCTGGATATTGAAGAGTCGTTTACTCATAAGA
CTGTTTAATATCACAGTTTATGATTTTTGTGTTAAAAACTGGTATAAGTTACCTTTTGCAGTTAGATTGC
GTATCATAAATAATACAAAACCTAAAACAGCAAGTACTATAAAACGTAGAAGAAGTATTGTTATTGATTA
CAGAAAAATTGCCATTCTCAACGCGTCGCGAAAATGATTGGCGGACTGTTTCTTAACACTCTAAGTTTTA
TAGTTACTACTCAACATGTGGTTGTTAATAACACACCACATTTTAATACTGTAGTGCAACAACATCATGT
TGTTAGTGCTAATGTTAAAAGTTTTCATCTAGAGTTCAGCATAGCTGTTCTCTTTGTTTTATTTCTAGCT
TTGTACCGTAGTACAAACTTTAAAGCGTGTGTCGGTATCTTAATGTTTAAGATAGTATCAATGACACTTA
TAGGACCTATGCTTATAGCATTTGGTCACTATATAGATGGCATTGTGACAACAACTGTCTTAGCTTTAAG
ATTCATTTACTTAGCATACTTTTGGTATGTTAATAGTAAATTTGAATTTATTTTATATAACACAACGACA

CTAATGTTTGTACATGACAGAGCTGCACCGTTTATGAGAAGTTCTCACGGCTCTATTTATGTCACATTGT
ATGGAGGCATAAATTATATGTTCGTGAATGATCTTACGTTGCATTTTGTAGACCCTATGCTTGTAGGCAT
AGCTATACGTGGCCTAGTTCGTGCTGACCTAACAGTTGTTAGAGCAGTTGAACTTCTCAATGGTGATTTT
ATCTATATATTTTCACAGGAGGCCGTCGTAGGAGTTTACAATGCAGCCTTTTCTCAGGCGGTTATAAACG
AAATTGATTTAAAAGAAGAAGAAGAACGTGTCTATGATGTTTCCTAGGGCATTTACTATCATAGACGACC
ATGGTATGGTTGTTAGTGTCTTCTTCTGGCTCCTGTTGATAATTGTATTGATATTGTTTTCAATAGCATT
GCTAAATGTTATTAAATTGTGCATGGTATGTTGTAATTTGGGTAAGACTATTATAGTATTACCCGCTCGC
CATGCATATGATGCCTATAAGACCTTTATGCAAATTAAGGCATATAATCCCGACGAAGCATTCTTAGTTT
GAACTAAACAAAATGAAGTACATTTTGTTCTTACTTGCGTGCATAATTGCATGCGTTTATGGAGAACGTT
ATTGTGCCATGCAAAGTGCAACTTCTACAAGCTGCATCAATGGCACTGACCACAACTCATGTCAAACTTG
CTTCGAACGTGGTGATCTTATTTGGCATCTCGCTAACTGGAACTTCAGCTGGTCTGTAATAT
TGATTGTTTTTATAACAGTGTTACAACCCGGGCCTCGGTAATGGCGAATGGGACCCA

Fragment 6B

GTCGACTGATTGTT
TTTATAACAGTGTTACAATACGGCAGACCGCAATTTAGCTGGCTCGTTTATGGCATTAAAATGCTGATCA
TGTGGCTATTATGGCCTATTGTTCTAGCGCTTACGATTTTTAATGCGTACTCTGAGTACGAAGTTTCCAG
ATATGTAATGTTCGGCTTTAGTGTTGCAGGTGCAGTTGTAACGTTTGCATTATGGATGATGTATTTTGTG
AGATCTATTCAGCTATATAGAAGGACCAAATCATGGTGGTCTTTTAACCCTGAGACTAATGCAATTCTCT
GTGTTAATGCGTTGGGTAGAAGCTACGTCCTACCCCTTGATGGCACTCCTACGGGTGTCACTCTCACGTT
ACTTTCAGGAAACCTATATGCTGAAGGGTTTAAAATGGCTGGTGGTCTTACCATCGAGCACTTGCCTAAA
TACGTCATGATTGCTACACCTAGTAGAACCATCGTTTATACCTTAGTTGGAAAACAGTTAAAAGCAACTA
CTGCCACAGGATGGGCTTACTATGTAAAATCTAAAGCCGGTGATTACTCAACAGAAGCACGTACTGATAA
TTTGAGTGAACATGAAAAATTATTACATATGGTATAACTAAACTCCTAAATGGCCACACAGGGACAACGC
GTCAACTGGGGAGATGAACCTTCCAAAAGACGTGGTCGTTCTAACTCTCGTGGTCGGAAGAATAATGATA
TACCTTTATCATTCTTTAACCCCATCACCCTCGAAAAAGGATCAAAATTTTGGAATTTATGTCCGAGAGA
CTTTGTTCCCAAGGGAATAGGTAATAAAGATCAACAAATTGGTTATTGGAATAGACAAGCGCGTTACCGC
ATTGTCAAGGGTCAGCGTACTGATCTTCCAGAGAGATGGTTCTTCTATTTCTTAGGTACAGGACCTCACG
CTGATGCTAAATTCAAAGACAAGATTGATGGAGTCTACTGGGTTGCTAGGGATGGAGCCACGAACAAGCC
TACAACACTTGGCACTCGTGGAACCAATAATGAGTCCAAACCATTGAAGTTTGACGGTAAGATACCACCG
CAATTTCAACTTGAAGTAAACCGTTCTAGGAACAATTCAAGGTCTGGTTCTCAGTCTAGATCTGTCTCAA
GAAACAGGTCTCAATCCAGAGGAAGACAACAATCCAATAACCAGAATACTAATGTTGAAGATACAATTGT
TGCTGTGCTTCAAAAATTAGGTGTTACTGACAAGCAAAGGTCACGTTCTAAATCCAGTGAACGTAGTGAT
TCAAAACCTAGAGATACAACACCTAAAAATGCCAACAAACACACCTGGAAGAAAACTGCAGGTAAGGGAG
ATGTGACAAACTTCTATGGTGCTAGAAGTGCTTCAGCTAACTTTGGTGATAGTGATCTCGTTGCCAATGG
TAACGCTGCCAAATGCTACCCTCAGATAGCTGAATGCGTTCCATCAGTGTCTAGCATGCTCTTCGGTAGT
CAATGGTCTGCTGAAGAAGATGGAGATCAAGTGAAAGTCACGCTCACTCACACCTACTACCTGCC
GTCGAC

Fragment 6C

GTCGACCGCTCACTCACACCTACTACCTGCCAAAGG
ATGATGCCAAAACCAGCCAGTTCCTAGAACAGATTGATGCCTACAAACGGCCTTCACAGGTGGCTAAAGA
TCAGAAGACAAGGAAATCCCGACACAAGTCCGCTGATAAGAAGCCCGAGGAACTGTCTGTTACTCTTGTA
GAGGCATACACAGATGTGTTTGATGACACACAGGTTGAGATGATTGATGAGGTTACGAACTAAACGCATG
CTCGTTTTCCTCCATGCTGTACTTGTAACAGCTTTAATCTTACTACTAATTGGTAGAATCCAATTACTAG
AAAGGTTGTTACTCAGTCATCTGCTTAATCTCACAACAGTCAGTAATGTTTTAGGTGTGCCTGACAGTAG
TCTGCGTGTAAATTGCTTACAGCTTTTGAGACCAGACTGCCTTGATTTTAATATCTTACATAAAGTTTTA
GCAGAAACCAGATTACTAGTAGTAGTACTGCGAGTGATCTTTCTAGTTCTTCTAGGGTTTTCCTGCTATA
CATTGTTGGGTGCATTATTTTAACATCATGATTGCTGTATTCCTTGTGTGTATTCTTTTGGCTAATGGAA
TTAAAGCTATTGATGTGCAACACGACCTTCATGAACATCCAGTCCTAACTTGGGAATTATTGCAACATTT
CATAGGAAGTACTCTTTACATCACAACACATCAGGTCTTAGCATTACCGCTTGGATCTCGTGTTGAATGT
GAGAGTATCGAAGGTTTTAATTGCACATGGCCTGGCTTTCAAAATCCTGCACATGATCACATTGATTTTT
ATTTTGATCTTTCTAATCCTTTCTATTCCTTTGTAGATAATTTCTATATAGTAGGTGAAGGAAATCAAAG
AATTAATCTCAGATTAGTTGGTGCTGTGCCAAAACAAAAGAGATTAAATGTTGGTTGTCATACATCATTT

GCTGTTGACCTTCCTTTTGGAACTCAGATCTACCATGACAGGGATTTTCAACTACCTGTTGATGGTAGAC
ATCTAGAGTGTGCTCACAGAGTTTACTTTGTGAAGTATTGTCCACAAAACCTGCATGGTTATTGCTTTAA
TGAGAAGTTGAAAGTCTACGACTTGAAGCAACTCAGAAGCAAAAAGGTCTTCGACAAACTCAACCAACAT
CATAAAACTGAGTTGTAAGGCAACCCGATGTTTAAAACTGGTCTTTCCGAGGAATTACTGGTCATCGCGC
TGTCTACTCTTGTACAGAATGGTAAGCACGTGTAATAGGAGGTACAAGCAACCCTATTGCATATTAGGAA
GTTTAGATTTGATTTGGCAATGCTAGATTTAGTAATTTAGAGAAGTTTAAAGATCCGCTAT
GACGAGCCAACAATGGAAGAGCTAACGTCTGGATCTAGTAATTGTTTAAAATGTAAAATTGTTTGAAAAT
TTTCCTTTTGATAGTGATACACAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATTCCGAGG
GGACCGTCCCCTCGGTAATGGCGAATGGGACCCAGTCGAC

Fragment 5AT2

CGTCTCCCATGATTGTGCTCGTAACTTGCCTCT
TGTTGTTATGTTCATACCACACAGTTTTGAGTACAACAAATAATGAATGCATACAAGTTAACGTAACACA
ATTGGCTGGCAATGAAAACCTTATCAGAGATTTTCTGTTTAGTAATTTTAAAGAAGAAGGAAGTGTAGTT
GTTGGTGGTTATTACCCTACAGAGGTGTGGTACAACTGCTCTAGAACAGCTCGAACTACTGCCTTTCAGT
ATTTTAATAATATACATGCCTTTTATTTTGTTATGGAAGCCATGGAAAATAGCACTGGTAATGCACGTGG
TAAACCATTATTATTTCATGTGCATGGTGAGCCTGTTAGTGTTATTATATCGGCTTATAGGGATGATGTG
CAACAAAGGCCCCTTTTAAAACATGGGTTAGTGTGCATAACTAAAAATCGCCATATTAACTATGAACAAT
TCACCTCCAACCAGTGGAATTCCACATGTACGGGTGCTGACAGAAAAATTCCTTTCTCTGTCATACCCAC
GGACAATGGAACAAAAATCTATGGTCTTGAGTGGAATGATGACTTTGTTACAGCTTATATTAGTGGTCGT
TCTTATCACTTGAACATCAATACTAATTGGTTTAACAATGTCACACTTTTGTATTCACGCTCAAGCACTG
CTACCTGGGAATACAGTGCTGCATATGCTTACCAAGGTGTTTCTAACTTCACTTATTACAAGTTAAATAA
CACCAATGGTCTAAAAACCTATGAATTATGTGAAGATTATGAACATTGCACTGGCTATGCTACCAATGTA
TTTGCTCCGACATCAGGTGGTTACATACCTGATGGATTTAGTTTTAACAATTGGTTCTTGCTTACAAATA
GTTCCACTTTTGTTAGTGGCAGGTTTGTAACAAATCAACCATTATTGATTAATTGCTTGTGGCCAGTGCC
CAGTTTTGGTGTAGCAGCACAAGAATTTTGTTTTGAAGGTGCACAGTTTAGCCAATGTAATGGTGTGTCT
TTAAATAACACAGTGGATGTTATTAGATTCAACCTTAATTTCACTGCAGATGTACAATCTGGTATGGGTG
CTACAGTATTTTCACTGAATACAACAGGTGGTGTCATTCTTGAAATTTCATGTTATAGTGACACAGTGAG
TGAGTCTAGTTCTTACAGTTATGGTGAAATCCCGTTCGGCATAACTGACGGACCACGATACTGTTATGTA
CTTTACAATGGCACAGCTCTTAAATATTTAGGAACATTACCACCCAGTGTAAAGGAAATTGCTATTAGTA
AGTGGGGCCATTTTTATATTAATGGTTACAATTTCTTTAGCACATTTCCTATTGGTTGTATATCTTTTAA
TTTAACCACTGGTGTTAGTGGAGCTTTTTGGACAATTGCTTACACATCGTATACTGAAGCATTAGTACAA
GTTGAAAACACAGCTATTAAAAATGCCCGGGAAAAGTGCATGTCCATTAAGAGACG

Fragment 5BT2

GTCGAC
GTTGAAAACACAGCTATTAAAAATGTGACGTATTGTAACAGTCACATTAATAACATTAAATGTTCTCAAC
TTACTGCTAATTTGAATAATGGATTTTATCCTGTTGCTTCAAGTGAAGTAGGTTTCGTTAATAAGAGTGT
TGTGTTATTACCTAGCTTTTTCACATACACCGCTGTCAATATAACCATTGATCTTGGTATGAAGCTTAGT
GGTTATGGTCAACCCATAGCCTCGACACTAAGTAACATCACACTACCAATGCAGGATAACAATACTGATG
TGTACTGTATTCGTTCTAACCAATTCTCAGTTTATGTTCATTCCACTTGCAAAAGTTCTTTATGGGACAA
TATTTTTAATCAAGACTGCACGGATGTTTTAGAGGCTACGGCTGTTATAAAAACTGGTACTTGTCCTTTC
TCATTTGATAAATTGAACAATTACTTGACTTTTAACAAGTTCTGTTTGTCGTTGAGTCCTGTTGGTGCTA
ATTGCAAGTTTGATGTTGCTGCACGTACAAGAACCAATGAGCAGGTTGTTAGAAGTCTATATGTAATATA
TGAAGAAGGAGACAACATAGTGGGTGTACCGTCTGATAATAGCGGTCTGCACGATTTGTCTGTGCTACAC
CTAGACTCCTGTACAGATTACAATATATATGGTAGAACTGGTGTTGGTATTATTAGACGAACTAACAGTA
CGCTACTTAGTGGCTTATATTACACATCACTATCAGGTGATTTGTTAGGCTTTAAAAATGTTAGTGATGG
TGTCATTTATTCTGTGACGCCATGTGATGTAAGCGTACAAGCGGCTGTTATTGATGGTGCCATAGTTGGA
GCTATGACTTCCATTAACAGTGAACTGTTAGGTCTAACACATTGGACAACGACACCTAATTTTTATTACT
ACTCTATATATAATTATACAAGTGAGAGGACTCGTGGCACTGCAATTGACAGTAACGATGTTGATTGTGA
ACCTGTCATAACCTATTCTAATATAGGTGTTTGTAAAAATGGTGCTTTGGTTTTTATTAACGTCACACAT
TCTGACGGAGACGTGCAACCAATTAGCACTGGTAATGTCACGATACCTACAAATTTTACTATATCTGTGC
AAGTTGAATACATGCAGGTTTACACTACACCAGTATCAATAGATTGTGCAAGATACGTTTGTAATGGTAA
CCCTAGATGTAACAAATTGTTAACACAATATGTGTCTGCATGTCAAACTATTGAACAAGCACTTGCAATG

GGTGCCAGACTTGAAAACATGGAGGTTGATTCCATGTTGTTTGTCTCGGAAAATGCCCTTAAATTGGCAT
CTGTTGAGGCGTTCAATAGTACAGAAAATTTAGATCCTATTTACAAAGAATGGCCTAGCATAGGTGGTTC
TTGGCTAGGAGGTCTAAAAGATATACTACCGTCCCATAATAGCAAACGTAAGTATGGTTCTGCTATAGAA
GATTTGCTTTTTGATAAAGTTGTAACATCTGGTTTAGGTACAGTTGATGAAGATTATAAACGTTGTACTG
GTGGTTACGACATAGCAGACTTGGTGTCGAC

Fragment 5CT2

GTCGAC
GTGGTTACGACATAGCAGACTTGGTGTGTGCTCAATATTACAATGGCATCATGGTTCTACCAGGTGTAGC
TAATGCTGACAAGATGACTATGTACACAGCATCACTTGCAGGTGGTATAACATTAGGTGCACTTGGTGGT
GGCGCCGTGGCTATACCTTTTGCAGTAGCAGTACAGGCTAGACTTAATTATGTTGCTCTACAAACTGATG
TATTGAATAAAAACCAACAGATCCTGGCTAATGCTTTCAATCAAGCTATTGGTAACATTACACAGGCTTT
TGGTAAGGTTAATGATGCTATACATCAAACATCACAAGGTCTTGCCACTGTTGCTAAAGCGTTGGCAAAA
GTGCAAGATGTTGTCAACACACAAGGGCAAGCTTTAAGTCACCTTACAGTACAATTGCAAAATAATTTTC
AAGCCATTAGTAGTTCTATTAGTGATATTTATAACAGGCTTGACGAACTGAGTGCTGATGCACAAGTTGA
TAGGCTGATTACAGGTAGACTTACAGCACTTAATGCATTTGTGTCTCAGACTCTAACCAGACAAGCAGAG
GTTAGGGCTAGTAGACAACTTGCCAAAGACAAGGTTAATGAATGTGTTAGGTCTCAGTCTCAGAGATTCG
GATTCTGTGGTAATGGTACACATTTGTTTTCACTAGCAAATGCAGCACCAAATGGCATGATTTTCTTTCA
TACAGTACTATTACCAACAGCTTATGAAACTGTAACAGCTTGGTCAGGTATTTGTGCTTCAGATGGCGAT
CGCACTTTCGGACTTGTCGTTAAAGATGTGCAGTTGACGTTGTTTCGTAATCTAGATGACAAGTTCTATT
TGACCCCCAGAACTATGTATCAGCCTAGAGTTGCAACTAGTTCTGATTTTGTTCAAATTGAAGGGTGTGA
TGTGTTGTTTGTCAACGCGACTGTAATTGATTTGCCTAGTATTATACCTGACTATATTGACATTAATCAA
ACTGTTCAAGACATATTAGAAAATTACGAACCAAACTGGACTGTACCTGAATTTACACTTGATATTTTCA
ACGCAACCTATTTAAATCTGACTGGTGAAATTGATGACTTAGAGTTTAGGTCAGAAAAGCTACATAACAC
TACAGTAGAACTTGCCATTCTCATTGATAACATTAATAATACATTAGTCAATCTTGAATGGCTCAATAGA
ATTGAAACTTATGTAAAATGGCCTTGGTATGTGTGGCTACTGATAGGTTTAGTAGTAGTATTTTGCATAC
CATTACTGCTATTTTGCTGTTTTAGCACAGGTTGTTGTGGATGCATAGGTTGTTTAGGAAGTTGTTGTCA
CTCTATATGTAGTAGAAGACAATTTGAAAATTATGAACCAATTGAAAAAGTGCATGTCCATTAAGAGACG
GTCGAC

Fragment 5AR

CGTCTCCCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACG
TAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCA
TCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCA
GCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCA
CCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACC
GCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACA
GCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCG
AGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG
ACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGG
AGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTCCGGACTCAGATCTAGGAGACGAC
CTTCCATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCAGGGCCGTACGCACCCTCG
CCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGATCCGGACCGCCACATCGAGCGGGTCACCGAGC
TGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGGAAGGTGTGGGTCGCGGACGACGGCCCCGCGGTGG
CGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGA
GCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGT
TCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGG
AGGCGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGC
TCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTG
CCTAAATTAAGAGACGGTCGAC

Fragment 6AR

CGTCTCCATTAACTAAAC
GATTTATGGACACTGTCAAGTCCATTGGCATCTCTGTAGACGCTGTACTTGACGAGTTAGATTCTATTGC
TTTTGCTGTAACACTTAAAGTATTATTTACTTCTGGTAGATTACTTGTGTGTATAGGTTTTGGTGATACT
TTTGAGGAGGCTGAACAAAAAGCTTATGCCAAACTTCAGCTGGATATTGAAGAGTCGTTTACTCATAAGA
CTGTTTAATATCACAGTTTATGATTTTTGTGTTAAAAACTGGTATAAGTTACCTTTTGCAGTTAGATTGC
GTATCATAAATAATACAAAACCTAAAACAGCAAGTACTATAAAACGTAGAAGAAGTATTGTTATTGATTA
CAGAAAAATTGCCATTCTCAACGCGTCGCGAAAATGATTGGCGGACTGTTTCTTAACACTCTAAGTTTTA
TAGTTACTACTCAACATGTGGTTGTTAATAACACACCACATTTTAATACTGTAGTGCAACAACATCATGT
TGTTAGTGCTAATGTTAAAAGTTTTCATCTAGAGTTCAGCATAGCTGTTCTCTTTGTTTTATTTCTAGCT
TTGTACCGTAGTACAAACTTTAAAGCGTGTGTCGGTATCTTAATGTTTAAGATAGTATCAATGACACTTA
TAGGACCTATGCTTATAGCATTTGGTCACTATATAGATGGCATTGTGACAACAACTGTCTTAGCTTTAAG
ATTCATTTACTTAGCATACTTTTGGTATGTTAATAGTAAATTTGAATTTATTTTATATAACACAACGACA
CTAATGTTTGTACATGACAGAGCTGCACCGTTTATGAGAAGTTCTCACGGCTCTATTTATGTCACATTGT
ATGGAGGCATAAATTATATGTTCGTGAATGATCTTACGTTGCATTTTGTAGACCCTATGCTTGTAGGCAT
AGCTATACGTGGCCTAGTTCGTGCTGACCTAACAGTTGTTAGAGCAGTTGAACTTCTCAATGGTGATTTT
ATCTATATATTTTCACAGGAGGCCGTCGTAGGAGTTTACAATGCAGCCTTTTCTCAGGCGGTTATAAACG
AAATTGATTTAAAAGAAGAAGAAGAACGTGTCTATGATGTTTCCTAGGGCATTTACTATCATAGACGACC
ATGGTATGGTTGTTAGTGTCTTCTTCTGGCTCCTGTTGATAATTGTATTGATATTGTTTTCAATAGCATT
GCTAAATGTTATTAAATTGTGCATGGTATGTTGTAATTTGGGTAAGACTATTATAGTATTACCCGCTCGC
CATGCATATGATGCCTATAAGACCTTTATGCAAATTAAGGCATATAATCCCGACGAAGCATTCTTAGTTT
GAACTAAACAAAATGACTTCGAAAGTTTATGATCCAGAACAAAGGAAACGGATGATAACTGGTCCGCAGTGGTGG
GCCAGATGTAAACAAATGAATGTTCTTGATTCATTTATTAATTATTATGATTCAGAAAAACATGCAGAAAATGCT
GTTATTTTTTTACATGGTAACGCGGCCTCTTCTTATTTATGGCGACATGTTGTGCCACATATTGAGCCAGTAGCG
CGGTGTATTATACCAGACCTTATTGGTATGGGCAAATCCCGGGCCTCGGTAATGGCGAATGGGACCCA

Fragment 6BR

GTCGACCCAGACCTTATTGGTATGGGCAAATCAGGCAAATCTGGTAATGGTTCTTATAGGTTACTTGATCATTAC
AAATATCTTACTGCATGGTTTGAACTTCTTAATTTACCAAAGAAGATCATTTTTGTCGGCCATGATTGGGGTGCT
TGTTTGGCATTTCATTATAGCTATGAGCATCAAGATAAGATCAAAGCAATAGTTCACGCTGAAAGTGTAGTAGAT
GTGATTGAATCATGGGATGAATGGCCTGATATTGAAGAAGATATTGCGTTGATCAAATCTGAAGAAGGAGAAAAA
ATGGTTTTGGAGAATAACTTCTTCGTGGAAACCATGTTGCCATCAAAAATCATGAGAAAGTTAGAACCAGAAGAA
TTTGCAGCATATCTTGAACCATTCAAAGAGAAAGGTGAAGTTCGTCGTCCAACATTATCATGGCCTCGTGAAATC
CCGTTAGTAAAAGGTGGTAAACCTGACGTTGTACAAATTGTTAGGAATTATAATGCTTATCTACGTGCAAGTGAT
GATTTACCAAAAATGTTTATTGAATCGGACCCAGGATTCTTTTCCAATGCTATTGTTGAAGGTGCCAAGAAGTTT
CCTAATACTGAATTTGTCAAAGTAAAAGGTCTTCATTTTTCGCAAGAAGATGCACCTGATGAAATGGGAAAATAT
ATCAAATCGTTCGTTGAGCGAGTTCTCAAAAATGAACAATAATATAACTAAACTCCTAAATGGCCACACAGGGAC
AACGCGTCAACTGGGGAGATGAACCTTCCAAAAGACGTGGTCGTTCTAACTCTCGTGGTCGGAAGAATAATGATA
TACCTTTATCATTCTTTAACCCCATCACCCTCGAAAAAGGATCAAAATTTTGGAATTTATGTCCGAGAGA
CTTTGTTCCCAAGGGAATAGGTAATAAAGATCAACAAATTGGTTATTGGAATAGACAAGCGCGTTACCGC
ATTGTCAAGGGTCAGCGTACTGATCTTCCAGAGAGATGGTTCTTCTATTTCTTAGGTACAGGACCTCACG
CTGATGCTAAATTCAAAGACAAGATTGATGGAGTCTACTGGGTTGCTAGGGATGGAGCCACGAACAAGCC
TACAACACTTGGCACTCGTGGAACCAATAATGAGTCCAAACCATTGAAGTTTGACGGTAAGATACCACCG
CAATTTCAACTTGAAGTAAACCGTTCTAGGAACAATTCAAGGTCTGGTTCTCAGTCTAGATCTGTCTCAA
GAAACAGGTCTCAATCCAGAGGAAGACAACAATCCAATAACCAGAATACTAATGTTGAAGATACAATTGT
TGCTGTGCTTCAAAAATTAGGTGTTACTGACAAGCAAAGGTCACGTTCTAAATCCAGTGAACGTAGTGAT
TCAAAACCTAGAGATACAACACCTAAAAATGCCAACAAACACACCTGGAAGAAAACTGCAGGTAAGGGAG
ATGTGACAAACTTCTATGGTGCTAGAAGTGCTTCAGCTAACTTTGGTGATAGTGATCTCGTTGCCAATGG
TAACGCTGCCAAATGCTACCCTCAGATAGCTGAATGCGTTCCATCAGTGTCTAGCATGCTCTTCGGTAGT
CAATGGTCTGCTGAAGAAGATGGAGGTCGAC

Fragment 6CR

GTCGACCAATGGTCTGCTGAAGAAGATGGAGATCAAGTGAAAGTCACGCTCACTCACACCTACTACCTGC
CAAAGG
ATGATGCCAAAACCAGCCAGTTCCTAGAACAGATTGATGCCTACAAACGGCCTTCACAGGTGGCTAAAGA
TCAGAAGACAAGGAAATCCCGACACAAGTCCGCTGATAAGAAGCCCGAGGAACTGTCTGTTACTCTTGTA
GAGGCATACACAGATGTGTTTGATGACACACAGGTTGAGATGATTGATGAGGTTACGAACTAAACGCATG
CTCGTTTTCCTCCATGCTGTACTTGTAACAGCTTTAATCTTACTACTAATTGGTAGAATCCAATTACTAG
AAAGGTTGTTACTCAGTCATCTGCTTAATCTCACAACAGTCAGTAATGTTTTAGGTGTGCCTGACAGTAG
TCTGCGTGTAAATTGCTTACAGCTTTTGAGACCAGACTGCCTTGATTTTAATATCTTACATAAAGTTTTA
GCAGAAACCAGATTACTAGTAGTAGTACTGCGAGTGATCTTTCTAGTTCTTCTAGGGTTTTCCTGCTATA
CATTGTTGGGTGCATTATTTTAACATCATGATTGCTGTATTCCTTGTGTGTATTCTTTTGGCTAATGGAA
TTAAAGCTATTGATGTGCAACACGACCTTCATGAACATCCAGTCCTAACTTGGGAATTATTGCAACATTT
CATAGGAAGTACTCTTTACATCACAACACATCAGGTCTTAGCATTACCGCTTGGATCTCGTGTTGAATGT
GAGAGTATCGAAGGTTTTAATTGCACATGGCCTGGCTTTCAAAATCCTGCACATGATCACATTGATTTTT
ATTTTGATCTTTCTAATCCTTTCTATTCCTTTGTAGATAATTTCTATATAGTAGGTGAAGGAAATCAAAG
AATTAATCTCAGATTAGTTGGTGCTGTGCCAAAACAAAAGAGATTAAATGTTGGTTGTCATACATCATTT
GCTGTTGACCTTCCTTTTGGAACTCAGATCTACCATGACAGGGATTTTCAACTACCTGTTGATGGTAGAC
ATCTAGAGTGTGCTCACAGAGTTTACTTTGTGAAGTATTGTCCACAAAACCTGCATGGTTATTGCTTTAA
TGAGAAGTTGAAAGTCTACGACTTGAAGCAACTCAGAAGCAAAAAGGTCTTCGACAAACTCAACCAACAT
CATAAAACTGAGTTGTAAGGCAACCCGATGTTTAAAACTGGTCTTTCCGAGGAATTACTGGTCATCGCGC
TGTCTACTCTTGTACAGAATGGTAAGCACGTGTAATAGGAGGTACAAGCAACCCTATTGCATATTAGGAA
GTTTAGATTTGATTTGGCAATGCTAGATTTAGTAATTTAGAGAAGTTTAAAGATCCGCTAT
GACGAGCCAACAATGGAAGAGCTAACGTCTGGATCTAGTAATTGTTTAAAATGTAAAATTGTTTGAAAAT
TTTCCTTTTGATAGTGATACACAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATTCCGAGG
GGACCGTCCCCTCGGTAATGGCGAATGGGACCCAGTCGAC

Infectious clone final unit sequences

NNNN BsmBI
NNNN Mutations

Unit 1

ACTTGGGCCCTAATACGACTCACTATAGACTTTTAAAGTTAAGTGAGTGTAGCGTGGCTA
TAACTCTTCTTTTACTTTAACTAGCCTTGTGCTAGATTTGTCTTCGGACACCAACTCGAA
CTAAACGAAATATTTGTCTCTCTATGAAACCATAGAAGACAAGCGTTGATTATTTCACCA
GTTTGGCAATCACTCCTAGGAACGGGGTTGAGAGAACGGCGCACCAGGGTTCCGTCCCTG
TTTGGTAAGTCGTCTAGTATTAGCTGCGGCGGTTCCGCCCGTCGTAGTTGGGTAGACCGG
GTTCCGTCCTGTGATCTCCCTCGCCGGCCGCCAGGAGAATGAGTTCCAAACAATTTAAGA
TCCTCGTTAATGAGGACTACCAAGTCAACGTGCCTAGTCTTCCTTTCCGTAACGCTTTGC
AGGAAATTAAGTACTGCTACCGTAACGGTTTTGAGGGCTATGTTTTCGTACCTGAATACC
GTCGTGACCTAGTTGATTGCGATCGTAAGGATCACTACGTCATTGGTGTTCTGGGTAACG
GAATAAGTGATCTTAAACCTGTTCTCCTTACCGAACCTTCCGTCATGTTGCAAGGCTTTA
TTGTTAGAGCTGGCTGCAATGGCGTTCTTGAGGACTTTGACCTTAAAATCGCCCGTACTG
GAAATGGCGCCATATATGTAGATCAATACATGTGTGGTGCGGATGGAAAACCAGTTATTG
AAGGTGAGTTTAAGGATTATTTCGGTGATGAGGACGTTATCATCTATGAAGGAGAGGAGT
ACCATTGTGCCTGGCTAACAGTGCGCGATGAAAAACCTTTATGTCAACAAAATCTCCTCA
CAATTAGGGAAATCCAATACAATCTGGACATCCCACATAAGTTGCCTAATTGTGCTATCA
GAGAGGTAGCACCACCAGTCAAGAAGAATTCCAAGGTTGTTCTTTCTGAAGAGTATAAAA
AGCTCTATGACATCTTTGGTTCACCATTTATGGGCAATGGGGACAGTCTTAACAAGTGTT
TTGACAGCCTCCACTTCATTGCTGCCACTCTTAAGTGTCCTTGTGGCGCCGAGAGTAGTG
GTGTTGGTGATTGGACTGGTTTTAAGACTGCTTGTTGTGGTCTTCAAGGCAAAGTTAAAG
GCGTCACATTAGGTGCTGTTAAACCAGGTGATGCCGTTGTCACTAGCATGAGTGCAGGTA
AGGGTGTAAAATTCTTTGCTAATAGTGTACTCCAATATGCTGGTGATGTGGAAAATGTCT
CTGTTTGGAAAGTGATTAAAACTTTCACTGTTAATGAGACTGTCTGCACCGCTGATTTTG
AAGGCGAATTGAACGACTTCATTAAACCAGAGAGTACTTCACCTGTTTCATGTTCTATTA
AGAGAGCGTTCATCACTGGTGAGGTTGATGATGCAGTCCATGACTGTATTATCTCAGGAA
AGTTGGACCTTAGTACTAACCTTTTTGGTGGTGCTAACCTGCTTTTCAAGAAAACGCCCT
GGTTTGTGTTAAAGTGTGGTGCTATTTTTGCTGATGCTTGGAAAGTTGTTGAAGAACTGC
TTTGTTCTCTCAAACTTACATACAAGCAAATCTATGATGTAGTAGCATCACTTTGCACTT
CTGCCTTTACTATTATGGATTACAAACCTGTATTTGTAGTTGCATCTAATAGTGTTAAGG
ATCTCGTTGACAAGTGTGTCAAAATTCTTGTGAAAGCCTTTGATGTTTTCACACAGACAA
TTACCATTGCGGGTGTGGAGGCTAAGTGCTTTGTGCTTGGCTCTAAATACTTGCTcTTTA
ATAATGCACTTGTTAAGCTTGTCAGCGTTAAAATACTTGGCAAGAGGCAGAAAGGTCTTG
ATAGTGCATTCTTTGCCACTAATTTGATTGGTGCTACTGTTAAGGTGACACCACAAAGAA
CTGAGGCAGCCAATATTAGTTTGAATAAAGTAGATGATGTTGTCACACCTGGTGAAGGTC
ATATTGTCATTGTAGGCGACATGGCTTTTTACAAGAGTGGAGAGTACTACTTTATGATGG
CAAGTCCTGATTCAGTTCTCATTAATAACGTTTTTAAAGCAGCTCGAGTGCCATCTTATA
ACATTGTTTACAATGTGGATGATGATACCAAGAGTAAGATGTCATCTAAAATTGGAACGC
CATTTGAATTTGATGGTGATGTGGATGCTGCTATTGTTAAGATTAACGACTTACTTGTTG
AGTTcAGACAAGAAAAGTTTTGTTTCAGAGCACTTAAGGATGGTGAAAGTATCTTGGTTG
AGGCATACCTTAAGAAATATAAAATGCCAATTTGTCTCAAAAATCATGTTGGTTTATGGG
ACATCATAAGACGCGATTCTGATAAGAAGGGTTTTCTTGACACTTTCAGTCATCTTAATG
AATTGGAGGATGTCAAGGATGTAAAAATTCAGTCTATTAAAAATATCCTTTGCCCTGACC
TTCTGTTAGAATTAGATTTTGGTGCAgTTTGGTATAGATGTATGCCAGCCTGCTCTGATG
TGTCTATTTTGGGTAATGTTAAAATCATGCTAGGTAATGGTGTTAAAGTGATTTGTGATG
GCTGCCATGGTTTTGCTAACCGTTTAACGGTTAATTACAGCAAACTGTGTGACACAGCTC
GTAAAGACATTGAAATTGGTGGTATACCATTTTCCACCTTTAAAACACCATCTAGCAGCT
TCATTGACATGAAAGATGCCATCTATTCTGTGGTTGAATACGGTGAAGCATTTTCCTTTA
AGACTGCAAGTGTACCTGTAACTAATAGTGGTACTGTCAAAAATGACGACTGGTCAGACC
CCATTTTGTTAGAACCTGCAGATTATGTCGAACCTAAAGATAACGGTGATGTAATTGTCA
TTGCTGGATACACCTTTTACAGGGATGAGAACGATCACTTTTATCCATATGGTTCCGGCA

TGGTTGTGCAGAAGATGTACAACAAGATGGGTGGTGGTGACAAGTCAGTTTCATTTTCAG
ACAACGTTAATGTTAAAGAGATAGAACCTATTACACGTGTCATGCTAGAATTTGAATTTG
ACAATGAGGTGGTAACACAGGTTCTTGAAAAAGTCATTGGTGCCAAATATAAGTTTACAG
GTACCACTTGGGAAGAATTTGAAGATTCTATTTCTGAGAAACTTGACAAGGTTTTCGACA
CGCTTGCTGAACAGGGTGTAGAACTTGAAGGTTACTTCATTTATGACACTTGTGGCGGAT
TTGACATTAATAATCCAGATGGTGTTATGATTTCACAGTATGATCTTAACTCAGCTGTTG
ATGTCAAAAGTGATTCTGATGCAAGCGTGGAAGACACATCTTCGATTTCTGATAACGAAG
ATGTTGAACAAATAGAAAAAGAGAATGTGTCAGCTGCTGATGTTACAGTTGTAGATGTTG
AAGAATTTGTTGAGCAAGTTTCTCTTGTTGACACCGAAGAACAACTCCCCTCAGTAGAGG
AAAAAGTTGAAGGCTCTGCTAAAGATGACCCGTGGGCTGCTGCTGTTGATGAACAGGAAG
CTGAACAACCTAAACCTTCCTTAACTCCATTTAAGACAACAAATCTTAATGGTAGAATTA
TTCTTAAACAACAAGATAACAATTGTTGGATAAACGCTTGCTGTTACCAGTTGCAAGCTT
TTGACTTTTTCAACCATGAATTATGGGATGGTTTTAAGAAAGATGACGTTATGCCTTTCG
TAAACTTCTGTTATGCAGCTCTAACTTTGAAGCAGGGTGATTCAGGAGACTCAGAATACC
TTTTGGAAACAATGCTCAATGATTATAGTACAGCTAAAGTTGTTCTTAGTGCTAAGTGTG
GCTGTGGTGTTAAAGAAATAGTTTTAGAGCGAACTGTGTTTAAGCTCACACCTCTTAAGA
ATGAGTTTAAATATGGTGTCTGTGGTGACTGTAAACAAATTAACATGTGTAAGTTTGCAA
GCGTGGAAGGTTCAGGTGTTTTTGTTCATGATAGAATTGAAAAACGAACACCTGTGTCAC
ACTTCATTGTGACACCTATTATGCACGCAGTTTACACAGGTACAACACAGAGTGGTCATT
ACATGATTGAGGATTGCACTCATGATTATTGTGTGGATGGTATGGGAGTTAAACCCCGTA
AACAAAAGTTTTACACATCCACCTTGTTTCTTAATGCAAATGTAATGACAGCTGACTCTA
AAACTACAATCGAACCAACTGCACCTGTTAAGGATAAGCTTGTAGAAGAATGTCAAAGTC
CTAAAGACTTTATATTACCCTTTTACACAGCGGGGAAAGTGTCATTTTACCAAGGAGATT
TGGATGTTTTGATTAACCTTCTTGAACCTGACGTTGTTGTTAACGCTGCTAATGGTGATC
TTAGGCACATGGGTGGTGTTGCTAAAGCCATTGATGTGTTCACAGGTGGTAAGTTGACGA
AACGTTCTAAAGAATATCTTAGGACTAATAAAACAATAGCTCCTGGTAATGCTGTGCTTT
TTGAAAACGTACTTGAGCATCTTAGTGTATTGAATGCTGTCGGACCACGTAATGGTGACA
GTAGAGTTGAAGGTAAACTCTGTAACGTTTACAAAGCAATTGCCAAGTGTGACGGTAAGA
TATTAACACCACTTATTAGTGTTGGTATCTTCAAAGTCAAACTGGAAGTTTCATTGCAGT
GTCTACTTAAAACTGTCACAGATAGGGAATTAAATGTCTTTGTTTACACTGACCAAGAGA
GAATTACCATAGAAAATTTCTTCAATGGTACTATTCCTGTCAAGGTAACGGAAGACACTG
TGAATCAAAAGCGTGTGTCTGTGGCACTTGATAAAACTTATGGTGAGCAACTCAAAGGCA
CTGTTGTTATTAAAGACAAAGATGTTACTAACCAGTTGCCAAGTGTGTCAGATGTTGGTG
AAAAAGTCGTTAAGGCATTGGATGTTGATTGGAATGCTTACTATGGCTTTTCAAATGCAG
CGGCTTTTAGTGCCAGTTCCCATGATGCTTATATATTTGAAGTTGTAACACACAACAACT
TCATTGTTCATAAGCAGACTGATAATAATTGTTGGGTTAATGCAATTTGTTTAGCATTAC
AAAGGCTTGAGACG

Unit 2

CGTCTCCGGCTTAAACCAACATGGAAGTTTCCTGGTGTTAAAAGTTTGTGGGATGATTTT
CTTGTTCGTAAAACTGCTGGATTTGTCCACATGCTATACCACATTTCAGGCTTAAAGAAA
GGACAACCTGGTGATGCTGAGTTGACTTTACATAAGCTTGGTGAGTTGATGTTGAGTGAT
AGTGCTGTCACCATCACACATACAACTGCATGTGACAAGTGTGCAAAAACTGAAACTTTT
ACAGGCCCTGTGATAGCAGCACCTCTTCTGATCCATGGCACTGATGAAACCTGTGTTCAT
GGTGTAAGTGTTAATGTTAAAGTGACAAGTGTTAGAGGCACTGTGGCTATCACTTCCTTA
ATCGGTCCTGTTGTAGGAGATGTTATTGATGCAACTGGTTATATTTGTTATACTGGATTG
AACTCACGTGGTCATTATACCTATTATGATAATCTTAATAGTTTAATGGTTGATGCTGAA
AAGGCCTATCATTTTGAGAAAAACCTCTTACAGGTTACAACAGCCATTGCTAGTAATTTT
ATTGCCTACACACCTAAGAGTGAGATCTATCCTAAGGTTCAACCAGTTAAAGAATCACAT
ACAGCTCGTGTTTTTAGTGAGGTTGAAGAAATTCCTAAAAATGTAGTTCGTGAAACAAAA
TTGCTAGCTATTGAGAGTGGTGCAGATTACACACTCTCAACACTTGGTAAATATGCTGAT
GTTTTCTTTATGACaGGcGATAAAATTCTTAGGTTTCTCCTTGAAGTCTTTAAATATTTA
TTGGTTGTTTTTATGTGTCTTAGAAAATCTAAGATGCCTAAAGTTAAGGTAAAACCACCG
CATGTATTCAAAAACTTAGGTGCTAAAGTTAGGACATTGAATTACGTAAGACAATTGAAC
AAACCAGCGTTATGGCGTTACGTAAAACTAGTGTTACTTTTAATAGCATTGTACCACTTT

TTCTACTTGTTTGTTAGTATACCAGTAATGCACAAATTAGTGTGTAGTAGTAGTGTACAA
GCGTATAGTAATTCTAGTTTTGTGAAATCAGAAGTTTGTGGTAATTCTATTCTATGCAAG
GCGTGTCTGGCATCTTATGATGAATTGGCTGATTTTGATCATCTTCAGGTGTCATGGGAT
TATAAGTCGGATCCGCTATGGAATAGAGTTATACAATTGTCTTACTTCACATTCTTGGTT
GTTTTTGGCAATAATTATGTTAGATGTCTTCTTATGTATTTTATCTCTCAGTACCTTAAC
TTGTGGTTGTCTTATTTTGGTTATGTTAAGTATAGTTGGTTTTTGCATGTCATCAACTTT
GAATCAATTTCTGTTGAATTTGTGATCATAGTTGTAGTTTTTAAGGCAGTTCTCGCACTT
AAACACATCTTCTTTTCATGCAATAACCCATCATGTAAGACTTGTTCTAAGATTGCTAGG
CAGACACGTATTCCTATACAAGTTGTGGTTAATGGTTCAATGAAGACTGTGTATGTCCAC
GCTAATGGTACTGGTAAACTCTGCAAAAAGCATAATTTCTACTGTAAAAATTGTGACTCT
TACGGTTTTGACCACACATTCATCTGTGACGAGATTGTGCGTGACCTGGGTAACAGTATC
AAGCAAACTGTTTATGCCACTGATAGATCATACCAGGAAGTCACAAAGGTTGAATGCTCT
GATGGCTTTTACAGGTTTTATGTTGGTGATGAATTCACGGCATACGACTATGATGTAAAA
AACAAGAAGTATAGTAGCCAGGAAGTGCTTAAAACTATGTTCCTGCTTGATGACTTCATC
GTCTATAGTCCATCAGGGTCTTCCCTCGCTAGTGTTAGAAATGTCTGTGTTTATTTTTCA
CAACTGATTGGTAGGCCTATTAAAATTGTTAATAGTGAACTTCTTGAGGATTTATCTGTA
GACTTTAAGGGAGCACTTTTTAATGCTAAGAAAAATGTTATTAAAAACTCTTTTAATGTT
GATGTGTCAGAATGTAAAAACCTTGAAGAGTGTTATAAAGCATGCAACCTTGATGTGACA
TTTTCTACTTTTGAAATGGCTGTTAATAACGCTCATAGATTTGGTATATTGATTACAGAC
CGCTCCTTTAACAATTTTTGGCCATCAAAAATTAAGCCAGGTTCTTCTGGTGTCTCTGCT
ATGGACATTGGTAAGTGCATGACTTTTGATGCAAAGATTGTTAATGCAAAGGTTTTAACG
CAACGAGGTAAAAGTGTTGTATGGCTTAGTCAGGACTTTTCTACTCTTAGTTCTACAGCA
CAGAAAGTTTTAGTTAAAACTTTTGTAGAAGAAGGTGTGAACTTTTCACTAACGTTTAAT
GCTGTGGGTTCAGATGAAGACCTCCCCTATGAGAGATTCACTGAATCAGTTTCACCAAAG
AAAGGTTCAGGTTTCTTTGATGTACTTAAACAGCTTAAACAGATCTTCTGGTGTTTGGTT
TTGTTTATTACCTTGTATGGTTTATGCTCAGTCTACAGTGTTGCAACTCAGTCTTATATT
GATTCTGCTGATGGTTATGAATACATGGTTATTAAGAATGGCATTGTTCAACCATTTGAT
GATTCTATTAACTGTGTACATAACACATACAAAGGCTTTGTTATGTGGTTTAAAGCTAAA
TATGGTTTTGTACCTACATTCGACAAGTCTTGTCCTATTGTTTTAGGAACTGTTTTTGAC
TTAGGCAATATGAGACCAATCCCTGATGTGCCAGCTTATGTAGCACTTGTTGGTAGATCC
CTAGTGTTTGCGATTAATGCAGCATTTGGTGTTACTAATGTTTGTTATGATCATACAGGT
TCTGCTGTGAGTAAAAACTCTTATTTTGATACCTGTGTATTTAATGCTGCTTGCACCACT
CTATCAGGCCTTGGAGGTACAATTGTTTATTGTGCCAAACAAGGACTTGTGGATGGTGCT
AAACTTTATAGCGAATTGATGCCTGACTATTTTTATGAGCACGCTAGTGGTAACATGGTT
AAAATACCAGCCATTATTAGGGGTTTTGGTCTGAGATTCGTAAAAACACAGGCTACAACT
TATTGCAGAGTAGGGGAGTGTGTTGAAAGTCAGGCCGGTTTTTGCTTTGGTGGTGACAAC
TGGTTTGTTTATGATAAAGAGTTTGGTGATGGTTATATTTGTGGAAGCTCTACATTGGGT
TTTCTCAAAAATGTCTTTGCTCTCTTTAACTCTAACATGTCTGTAGTAGCCACATCCGGT
GCTATGTTGGTTAACATTATCATCGCATGTTTGGCTATAGCGGTCTGTTATGGTGTTCTT
AAATTTAAGAAAATCTTTGGTGATTGTACTCTCCTAGTTGTTATGATTATTGTTACATTG
GTTGTTAACAATGTTTCTTATTTTGTGACCCAAAACACATTCTTTATGATTGTTTATGCC
ATTATTTATTACTTTACAACAAGGAAATTATCATACCCTGGTATCTTGGATGCAGGATTC
ATTATTGCCTATGTGAACATGGCCCCATGGTATGTGCTTGTGCTGTACATAGTGGTCTTC
TTATATGATTCTCTACCCTCACTATTCAAACTTAAGGTTACAACAAACCTTTTTGAAGGG
GATAAATTTGTTGGTAGTTTTGAATCAGCAGCTATGGGTACTTTTGTTATAGATATGCGT
TCATATGAGACACTTGTTAATTCTACATCCTTGGATAGAATTAAGTCATATGCTAACAGC
TTTAACAAGTACAAGTACTATACTGGTTCAATGGGAGAGGCCGATTACCGCATGGCTTGT
TATGCTCATTTAGGTAAGGCATTAATGGATTATTCAGTCTCAAGAAATGATGTTCTTTAC
ACACCACCAACGGTCAGTGTTAATTCAACATTACAGTCAGGACTGAGAAAAATGGCACAG
CCTAGTGGTATTGTGGAACCCTGCATTGTGAGGGTAGCCTATGGCAATAATGTGCTTAAT
GGTTTGTGGCTTGGAGATGAAGTTATCTGCCCTAGACACGTCATTGCAAGTGACACATCA
CGTGTGATCAATTATGACAATGAGTTGTCCGGTGTGCGGTTACATAACTTCTCTATAGCC
AAAAATAATGTGTTTTTGGGCGTTGTGTCTGCCAAGTATAAGGGTGTGAACCTTGTGCTT
AAGGTTAATCAGGTAAACCCCAACACACCAGAACACAAATTCAAATCCGTGAAGTCTGGT
GAGAGTTTTAACATTCTCGCTTGTTATGAAGGCTGTCCCGGTAGTGTCTACGGTGTGAAC
ATGAGAAGTCAGGGTACCATCAAAGGTTCATTTATTGCTGGTACCTGTGGGTCAGTAGGT
TATGTATTAGAAAATGGAATACTCTATTTTGTGTACATGCACCACTTAGAATTAGGTAAT

GGCTCTCATGTTGGTTCAAACCTTGAGGGGGTAATGTATGGCGGTTATGAAGATCAACCT
AGCATGCAATTGGAGGGTACTAACGTCATGTCATCAGATAATGTAGTTGCATTTTTGTAT
GCTGCTCTTATTAATGGTGAAAGATGGTTTGTTACAAATGCATCAATGACGTTAGAATCT
TACAATGTATGGGCCAAAACTAACAGTTTTACGGAACTTGTTTCAACTGATGCTTTTAAT
ATGTTGGCTGCAAAAACTGGTTATAGTGTTGAGAAGTTGCTTGAATGTATTGTTAGACTC
AACAAAGGTTTTGGAGGACGCACTATACTCTCTTATGGTTCATTGTGTGATGAATTCACA
CCTACGGAAGTTATAAGGCAAATGTATGGTGTTAATCTTCAGAGTGGTAAAGTGAAATCT
CTCTTTTACCCTGTGTTGACAGCCATGGCTATTCTATTTGCTTTTTGGTTAGAGTTCTTT
ATGTACACACCCTTTACATGGATTAATCCATCGTTTATCAGTGTTATTCTAGCTATTACA
ACTTTTTTCTCTGTGTTATTAGTTGTTGGGATTAAACACAAAATGTTATTCTTTATGTCT
TTTGTAATGCCTAGTGTTGTGTTGGCCACCGTTCATAATGTTGTATGGGATATGACTTAT
TATGAAAGCTTACAAGCTCTTGTGGCGAACACTAACACAACATTTTTACCAGTAGATATG
CAAGGCGAGACG

Unit 3

CGTCTCCAAGGCGTTATGCTTGCTTTGTTCTGTGTTGTTGTGTTTGTTACATACACTATA
AGATTCTTTACATGCAAGCAATCCTGGTTTTCTTTATTTGTAACAACTGTCTTTGTGGTG
TTCAACATTGTTAAGTTGCTAGGCATGGTAGGTGAACCATGGACTGAAGATCACATCTTG
TTATGTCTTGTAAACATGCTTACCATGTTGATAAGTCTTACTACAAAGGACTGGTTTGTT
GTTTTTGTATCATACAAATTTGCTTATTATGTCGTTGTTTATGTAATGCAACCATCCTTT
GTTCAAGACTTTGGGTTTATTAAGTGTGTAAGCGTGATTTACATGGCATGTGGTTATTTG
TTTTGCTGTTATTATGGTATTCTTTACTGGGTTAACAGATTTACATGCATGACGTGTGGT
GTTTACCAATTCACTGTCTCACCGGCAGAGCTCAAGTATATGACTGCTAACAATTTGTCA
GCACCTAAGACCGCATATGATGCCATGATTCTTAGTGTTAAATTGATGGGTATAGGTGGA
GAGAGAAATATAAAAATCTCCACTGTACAGTCTAAACTCACAGAGATGAAATGCACAAAC
GTTGTGCTGCTTGGGCTTTTGTCTAAAATGCATGTTGAATCTAACTCAAAGGAGTGGAAT
TATTGTGTTGGGCTGCATAATGAGATTAATTTGTGTGATGACCCCGAAGTGGTTCTTGAA
AAACTACTTGCACTCATCGCATTCTTTTTATCCAAACACAATACATGTGATCTTAGTGAC
CTTATCGACTCCTATTTTGATAACACCACTATTTTACAGAGTGTGGCTTCTGCTTATGCT
GCTTTACCTAGCTGGATTGCATATGAAAAAGCTCGCGCGGATTTAGAAGAGGCTAAGAAA
AATGATGTTAGTCCACAATTATTGAAACAGCTTACAAAAGCATGCAATATTGCTAAGAGT
GAGTTTGAACGCGAGTCCTCGGTTCAAAAGAAGCTTGATAAAATGGCCGAGCAAGCTGCA
GCTAGTATGTACAAAGAGGCAAGGGCTGTTGACAGAAAATCAAAAATAGTTTCTGCTATG
CATAGTCTTCTTTTTGGTATGCTCAAGAAACTTGATATGTCCAGTGTTAATACTATTATT
GAACAGGCCCGTAATGGTGTGTTGCCCCTGAGTATCATACCTGCTGCAGCAGCCACAAGA
CTTATTGTTGTAACACCTAATATTGAAGTGCTTTCCAAAATTAGGCAAGAAAACAATGTC
CACTATGCAGGTGCTATCTGGTCAATTATTGAAGTCAAGGATGCTAATGGTGCTCAAGTC
CATTTAAAAGAAGTCACAGCTGCCAATGAGTTGAACCTGACCTGGCCTTTGAGTATTACA
TGTGAGAGAACCACAAAGTTGCAAAACAATGAAATTATGCCAGGTAAACTTAAAGAGAAG
GCTGTTAAAGCGTCTGCTACAATTGATGGTGATGCTTATGGTAGTGGTAAGGCCCTCATG
GCTTCTGAAAGTGGTAAGAGCTTTATATATGCATTTATAGCGTCTGACAGCAACCTTAAG
TATGTTAAATGGGAGAGCAATAATGATGTAATACCCATTGAACTTGAGGCCCCATTGCGC
TTTTATGTAGATGGTGCTAATGGCCCTGAAGTTAAGTATCTTTACTTTGTGAAGAACTTG
AACACTCTTAGACGTGGTGCTGTTCTGGGTTATATTGGCGCTACTGTTCGTTTGCAGGCT
GGTAAACCTACTGAACACCCATCTAACAGTGGTTTATTAACCTTGTGTGCTTTCGCTCCT
GACCCTGCTAAGGCATACGTTGATGCTGTTAAGAGAGGTATGCAGCCAGTTACTAATTGT
GTGAAAATGCTCTCAAATGGTGCTGGTAATGGTATGGCTATTACAAATGGTGTAGAGTCT
AATACGCACCAAGACTCTTATGGTGGTGCATCTGTGTGCATCTATTGTAGGTGTCATGTA
GAACATCCTGCAATTGATGGATTGTGTCGTTATAAGGGTAAATTTGTTCAAGTACCAACC
GGCACTCAAGATCCTATTAGATTTTGTATTGAAAATGAAGTCTGTGTTGTTTGTGGTTGC
TGGCTTAATAATGGTTGCATGTGTGATCGTACTTCTATGCAAGGCACAACTGTTGACCAG
AGTTATTTAAACGAGTGCGGGGTTCTAGTGCAGCTCGACTAGAACCCTGTAATGGTACTG
ATCCAGACCACGTTAGTAGAGCTTTTGACATCTATAATAAAGATGTTGCATGTATTGGTA
AATTCCTTAAGACTAATTGTTCCCGTTTTAGGAATTTGGACAAACGTGATGCCTACTATG
TTGTCAAGCGTTGTACTAAGAGCGTTATGGACCATGAGCAAGTCTGTTATAACGATCTTA

AAGATTCTGGTGCAGTTGCTGAACATGACTTCTTTTTGTATAAGGAAGGTCGATGTGAAT
TCGGCAACGTCGCACGTAAGGATCTTACAAAGTACACTATGATGGATCTATGTTATGCCA
TTAGGAATTTTGATGAAAAGAATTGCGAAGTTCTTAAAGAAATACTCGTGACATTAGGTG
CCTGCAATGAATCCTTCTTTGAAAATAAAGATTGGTTTGACCCAGTGGAAAATGAAGCTA
TACATGAGGTGTACGCAAGACTTGGACCTATAGTAGCCAACGCTATGCTTAAATGTGTAG
CGTTTTGTGATGCTATCGTGGAGAAAGGTTACATAGGTATTATTACACTTGATAACCAAG
ACCTTAATGGTAATTTTTACGACTTCGGTGATTTCGTAAAGACAGCGCCTGGTTTTGGAT
GTGCTTGCGTGACGTCATATTATTCTTATATGATGCCCTTAATGGGAATGACTTCATGTT
TAGAGTCTGAAAATTTTGTGAAAAGTGACATTTATGGTTCTGATTATAAGCAGTATGATT
TGCTAGCTTATGATTTTACAGACCATAAAGAAAAGCTGTTCGAAAAATATTTCAAACACT
GGGACCGCACATACCATCCTAACTGCTCTGATTGTACAAGTGATGATTGTATCATTCATT
GTGCTAATTTTAACACATTGTTTTCAATGACCATACCCAACACTGCTTTTGGACCTCTTG
TACGTAAAGTCCACATAGATGGTGTGCCAGTTGTTGTTACTGCTGGTTACCATTTTAAAC
AGTTGGGTATTGTGTGGAATCTTGATGTTAAATTAGACACTATGAAGTTGACTATGACAG
ATCTTCTTAGGTTCGTCACTGACCCTACATTACTTGTAGCATCAAGCCCTGCATTGTTAG
ATCAGCGCACTGTCTGTTTTTCCATTGCAGCTTTGAGTACTGGTGTTACATATCAGACAG
TGAAACCAGGTCACTTTAACAAGGATTTCTACGATTTTATAACAGAGCGTGGATTCTTTG
AGGAGGGATCTGAGTTGACGTTGAAACATTTCTTCTTTGCACAGGGTGGTGAAGCCGCTA
TGACAGATTTTAACTACTACCGTTATAACAGAGTTACTGTTTTAGACATCTGTCAGGCGC
AATTTGTTTATAAAATTGTGTGTAAGTACTTTGATTGTTACGATGGTGGCTGCATCAATG
CTCGAGAGGTTGTTGTTACCAACTACGACAAGAGTGCTGGTTATCCACTTAATAAATTTG
GTAAGGCAAGACTTTACTATGAGACTCTGTCATATGAGGAGCAGGATGCGATTTTTGCAT
TAACAAAGAGAAATGTTCTACCCACAATGACACAAATGAATTTGAAATATGCTATATCTG
GAAAGGCTAGAGCTCGCACTGTAGGAGGAGTGTCACTCCTTTCGACAATGACTACGAGAC
AATACCACCAGAAGCATTTAAAGTCTATTGCTGCTACACGTAATGCCACTGTTGTTATTG
GAACGACTAAGTTCTACGGTGGCTGGGATAATATGTTAAAGAATTTGATGCGTGATGTAG
ATAACGGCTGTTTGATGGGATGGGACTATCCAAAGTGCGACCGTGCTTTGCCTAATATGA
TACGAATGGCTTCTGCAATGGTATTGGGTTCTAAGCACATTGGGTGTTGTACACATAGTG
ATAGGTTCTATCGCCTCTCTAACGAGTTAGCTCAAGTACTTACAGAGGTTGTGCATTGCA
CAGGTGGTTTTTATATCAAACCTGGTGGTACAACTAGTGGTGATGGTACTACAGCATATG
CTAATTCGGCTTTTAACATCTTTCAAGCTGTTTCTGCTAATGTTAATAAGCTTTTGGGAG
TTGACTCAAACACCTGTAACAATGTTACAGTTAAGTCCATACAACGCAAGATATATGATA
ATTGTTATCGAAGTAGCAGTGTTGATGATGACTTCGTTGTTGAGTATTTTAGCTATTTAA
GAAAGCACTTTTCTATGATGATTTTGTCAGATGATGGCGTGGTTTGTTACAACAAAGACT
ATGCTGATCTTGGTTATGTAGCAGACATTAGTGCTTTTAAAGCTACTTTGTATTACCAGA
ATAATGTTTTTATGTCTACAGCCAAGTGTTGGGTAGAACCAGACCTTAATGTTGGACCAC
ACGAATTTTGTTCACAACACACACTACAAATTGTAGGACCGGATGGTGATTATTATTTAC
CATACCCAGACCCTTCTAGAATCTTGTCAGCAGGTGTTTTCGTTGATGACATCGTTAAAA
CAGATAACGTCATTATGCTTGAACGTTATGTGTCTTTAGCAATCGATGCGTATCCACTTA
CAAAACACCCTAAACCCGCATATCAAAAAGTGTTTTATGCATTGTTAGATTGGGTTAAAC
ATTTACAGAAGACTCTAAATGCTGGTATACTTGATTCATTTTCTGTCACTATGTTGGAGG
ATGGTCAAGATAAATTCTGGAGTGAAGAGTTCTATGCCAGTCTTTACGAAAAGTCTACTG
TTTTGCAAGCTGCCGGTATGTGTATTGTTTGTGGTTCACAAACTGTGTTACGTTGCGGAG
ACTGTTTAAGGAGACCTCTTTTATGCACTAAGTGCGCTTACGACCATGTCATGGGTACAA
AGCATAAATTCATTATGTCTATCACACCATATGTGTGTAGCTATAATGGTTGTACTGTTA
ATGATGGAGACG

Unit 4

CGTCTCCTGATGTAACAAAATTGTTTTTGGGAGGTCTTAACTATTATTGTACTGAACATA
AACCACAATTGTCATTCCCGCTCTGTGCTAATGGCAATGTGTTTGGTCTGTACAAGAGTA
GTGCACTTGGTTCTGAGGACGTTGACGATTTCAATAAACTTGCAATCTCGGACTGGACCA
ATGTAGAAGATTACAAACTTGCTAATAATGTCAAAGAATCTCTGAAAATTTTCGCTGCTG
AAACTGTGAAAGCAAAGGAGGAGTCTGTTAAAGCTGAATATGCTTATGCCATATTAAAGG
AGGTGGTCGGCCCTAAGGAAGTTGTACTCCAATGGGAATCCTCTAAGGTTAAACCTCCAC
TTAACAGAAATTCTGTTTTCACATGTTTTCAGCTAAACAAGGATACTAAAATTCAGTTAG

GTGAATTTGTGTTTGAGCAGTCAGAGTATGGTAGTGAATCTGTTTACTACAAGAGTACTA
GCACTACTAAGCTGACACCGGGTATGATTTTTGTGTTGACATCTCATAACGTGAGTCCAC
TTAAAGCTAATATCTTAGTCAACCAAGAGAAGTACAATACCATATCCAAGCTCTATCCTA
CGTTCAACATAGCGGAAGCCTATACCACATTGGTGCCTTACTATCAAATGATTGGTAAGC
AAAAGTTTACAACTATCCAAGGTCCTCCTGGTAGTGGTAAATCACATTGTGTTATAGGTT
TGGGTTTGTATTATCCTCAAGCTAGGATCGTCTACACAGCATGTTCACATGCTGCTGTGG
ATGCATTGTGTGAAAAGGCTTCTAAGAACTTTAATGTGGATAAGTGTTCAAGGATAATAC
CCCAAAGAATCAGAGTTGACTGCTATATGGGTTTTAAACCTAATAATACTAATGCACAGT
ACTTGTTTTGTACAGTTAATGCTTTACCTGAAGCTAATTGCGACATCGTAGTTGTGGATG
AAGTCTCAATGTGTACAAATTATGATCTCAGTGTTATAAATAGTAGACTGAGTTACAAGC
ACATTGTTTACGTGGGAGACCCACAACAGCTTCCAGCTCCTAGAACCTTGATTAATAAGG
GTACACTCCAGCCTGAAGATTACAATGTTGTAACCCAGAGAATGTGTAAACTAGGACCTG
ATGTATTTTTGCATAAATGCTATAGATGCCCAGCTGAAATTGTTAAGACAGTCTCCGCGC
TTGTTTATGAGAATAAGTTTTTGCCTGTCAACCCAGAGTCAAAGCAGTGCTTTAAAATGT
TTGTAAAAGGTCAGGTGCAGATTGAATCCAACTCTTCTATAAACAATAAGCAACTAGATG
TTGTTAAGGCATTCCTAGCACATAATCCAAAATGGCGTAAAGCCGTTTTTATCTCACCCT
ATAACAGTCAGAACTATGTCGCAAAACGTCTACTCGGTTTGCAAACACAAACGGTAGACT
CTGCTCAAGGTAGTGAGTATGATTATGTCATTTACACACAGACTTCTGATACGCAGCATG
CTATTAATGTCAACAGGTTTAATGTTGCCATTACTAGAGCAAAGGTTGGTATTCTCTGTA
TCATGTGTGATAGAAAGATGTATGACAATCTTGATTTCTATGAACTCAAAGATTCAAAAG
TTGGCTTGCAAGCAAAACCTGAAACGTGTGGTTTGTTCAAAGATTGTTCAAAGACTGATC
AATACATACCACCAGCATATGCTACAACATATATGAGTCTGTCTGAAAATTTTAAGACAA
GTGATGGTTTAGCTGTTAACATCGGTACAAAGGATGTTAAATATGCTAATGTTATCTCAT
ATATGGGATTCAGGTTTGAGGCCAATGTACCAGGCTATCACACATTGTTTTGCACACGAG
ACTTTGCTATGCGTAATGTGAGATCATGGTTAGGGTTTGATGTCGAAGGTGCACATGTCT
GTGGTGATAACATCGGAACTAATGTACCACTACAGCTGGGTTTTTCAAATGGTGTAGATT
TTGTAGTACAAACTGAGGGATGTGTTGTTACTGAGAAAGGTAATAGTATTGAGGTTGTAA
AGGCAAGAGCACCACCGGGTGAGCAATTTGCACATTTGATTCCACTTATGAGAAAAGGTC
AGTCTTGGCACATTGTTAGACGTCGTATAGTGCAGATGGTTTGTGACTATTTCGACGGTT
TATCTGACATCCTAATTTTCGTGCTTTGGGCTGGTGGTCTTGAGCTTACAACTATGAGAT
ACTTTGTTAAAATTGGAAAACCACAGAAATGTGATTGCGGCAAAATTGCAACGTGTTATA
ACAGCTCGCAGTGTGTCTATGCTTGTTTTAAACATGCATTAGGATGTGATTACTTGTACA
ATCCTTACTGCATTGATATTCAACAATGGGGTTATACAGGATCCTTGAGCATGAACCATC
ATGAAGTTTGTAACATTCATAGAAATGAGCATGTAGCTAGTGGTGATGCTATCATGACTA
GGTGCCTTGCTATATATGATTGTTTTGTCAAGCGTGTAGATTGGTCCATTGTGTACCCAT
TTATTGATAATGAAGAAAAGATCAATAAAGCTGGTCGCATTGTACAATCACATGTCATGA
GAACTGCTCTTAAGGTTTTCAACCCTGCTGCAATTCACGATGTTGGTAATCCTAAAGGTA
TTCGTTGTGCTACAACATCTATACCATGGTTCTGTTATGACCGTGATCCTATTAATAACA
ATGTTAGATGTCTGGAGTATGATTACATGGTACATGGACAAATGAATGGTTTAATGTTAT
TTTGGAATTGCAACGTAGACATGTACCCGGAGTTCTCGATTGTTTGTAGATTTGATACTC
GAACTCGCTCAAAATTGTCATTAGAAGGTTGTAATGGTGGTGCATTGTATGTTAATAACC
ATGCATTTCACACACCAGCTTATGATAGAAGAGCATTTGCTAAGCTTAAACCTATGCCAT
TCTTTTACTACGATGAGAGTGATTGTGAGTTTATTGATGGACAACCTAATTATGTACCAC
TTAAAGCTAATGTCTGCATAACTAAATGCAACATAGGTGGTGCAGTCTGTAAGAAACATG
CAGCACTTTACAGAGCATATGTTGAAGACTACAACACTTTCATACAAGCTGGTTTTACAA
TTTGGTGCCCTCAAAATTTTGATACATATATGTTGTGGCAAGGTTTTGTTAATAGCAAAG
CACTCCAGAGTTTGGAAAACGTAGCCTTCAATGTCGTCAAGAAAGGTGCCTACGCTGATT
TAAAAGGTGACTTACCAACAGCAGTTATAGCTGATAAGATCATGGTGAGGGATGGACCTA
CTGACAAGTGTATTTTCACAAATAAAACTAGTTTGCCTACGAATGTGGCTTTTGAGCTGT
ATGCAAAACGCAAACTTGGACTCACACCTCCGTTAACAATACTTAGAAATCTAGGTGTTG
TCGCAACACATAAATTCGTGTTGTGGGATTATGAAGCCGAATGTCCTTTCTCAAACTTTA
CTAAGCAAGTGTGTGCGTACACTGATCTTGATGGTGAAGTTGTAACATGTTTTGATAATA
GTATTAGTGGTTCTTTTGAACGCTTTACTACTACGAAAGATGCAGTGCTTATTTCTAATA
ACGCTGTGAGGGGACTCAGTGCCATTAAATTACAATATGGCTTTTTGAATGATTTACCTG
TAAGTACTGTAGGAAACAAACCTGTCACATGGTACATCTATGTGCGCAAGAACGGTGAGT
ACGTCGAACAGATTGACAGTTATTACACACAGGGACGTACTTTTGAAACCTTTAAACCTC
GTAGTACAATGGAAGAAGACTTTTTTAGTATGGATACTACACTCTTCATCCAAAAGTATG

GTCTTGAGGATTATGGTTTTGAACATGTTGTATTCGGAGATGTTTCTAAAACTACCATTG
GTGGTATGCATCTTCTTATATCACAAGTGCGCCTAGCAAAAATGGGTTTGTTTTCTGTCC
AAGAATTTATGAGTAATTCTGACAGTACACTGAAAAGTTGTTGTATAACATACGCTGACG
ATCCGTCTTCTAAGAATGTGTGCACTTATATGGACATACTCTTAGACGATTTTGTGACTA
TCGTTAAGGGCTTAGATCTTAACGTTGTGTCGAAAGTTGTGGATGTTATTGTAGATTGTA
AGGCATGGAGATGGATGTTGTGGTGTGAGAACTCACAAATTAAAACCTTCTATCCACAAC
TCCAGTCTGCTGAGTGGAATCCTGGCTATAGCATGCCTACACTTTACAAAATACAGCGTA
TGTGTCTCGAACGGTGTAATCTCTACAATTATGGTGCACAAGTGAAGTTACCTGACGGCA
TTACTACTAATGTCGTTAAGTATACCCAGTTGTGTCAATATCTCAATACCACTACACTGT
GTGTTCCACATAAAATGCGTGTTCTGCATTTGGGAGCAGCAGGTGCTAATGGTGTTGCCC
CTGGTACCACTGTATTAAGAAGATGGTTGCCAGACGATGCCATATTGGTTGATAATGATA
TGAGGGATTACGTTTCCGACGCAGACTTCAGTGTTACAGGTGACTGTACTAATCTTTATA
TAGAGGACAAATTTGATTTACTTATCTCTGATTTATATGATGGCTCAACCAAGTCTATAG
ACGGTGAAAACATGTCGAAAGATGGCTTCTTTACATATATTAATGGTTTTATTAATGAGA
AGCTAGCGCTTGGAGGTTCTGTAGCCATCAAAATCACTGAATTTAGCTGGAATAAAAGTT
TATATGAATTAATTCAAAGATTTGAGTATTGGACTGTGTTTTGTACAAGTGTTAATACCT
CGTCGTCAGAAGGTTTTCTGATAGGTATCAACTACTTAGGACCATACTGTGACAAGGCTA
TAGTGGATGGAAATATAATGCATGCCAATTATATATTTTGGAGAAATTCCACAATTATGG
CGTTGTCACACAACTCAGTCCTAGATACTCCAAAATTCAAGTGTCGTTGCAATAACGCAC
TTGTTGTTAATTTAAAAGAAAGAGAATTAAATGACATGGTTATTGGATTACTAAGGAAGG
GTAAAATACTCATTAGAAATAATGGTAAGTTACTAAACTTTGGTAATCACTTTGTTAACA
CACCATGAGAGACG

Unit 5

CGTCTCCCATGATACTATTGTTGGTACTTTTTAGTGTTGTAGGAGCTCACGATGCACCTC
ATGGTGTAACACTACCACAGTTTAACACCTCATACAACAATGACAAGTTTGAACTTAATT
TCTACAATTTCTTGCAAACTTGGGATATACCACCAAACACAGAAACTATTCTGGGTGGTT
ACCTACCATACTGTGGGACAGGTGCTAACTGTGGTTGGTACAATTTTGTCTACCAACAAA
ATGTGGGATCAAATGGTAAATATTCTTATATAAACACACAAAATTTGAACATACCGAATG
TTCATGGTGTTTATTTTGACGTACGTGAGCACAATTCAGATGGTGTGTGGGATCTACGCG
ACCGTGTAGGTTTACTCATAGCCATCCACGGCAAATCACAATATAGCCTACTCATGGTTT
TGCAAGACAATGTGGAAGAGAATCAGCCTCACGTAGCTGTTAAAATCTGCCATTGGAAAC
CTGGCAACATAAGTTCAATTCATCAATTTAGTGTTAACCTAGGAGATGGTGGTCAATGTG
TGTTCAATCAGAGATTCTCATTGGACACCATATTGACAACTAATGACTTCTATGGCTTCC
AGTGGACTAATAATTATGTCAACATTTATCTAGGTGGCACCATTACTAAAGTGTGGGTTG
AAAATGATTGGAGTGTTGTTGAATCCAGTATTTCCTACCACTGGAACCGAATCAACTATG
GATATTATATGCAATTTGTCAACCGCACCACTTATTACGTTTATAATAATACTGGTGGTG
CAAATTATACTCACCTGCAGCTCGAAGAGTGCCACTCTGAGTATTGCGCTGGTTATGCTA
AGAATGTCTTTGTGCCTATTGATGGTAAAATACCAGATGGCTTCTCTTTTAGTAATTGGT
TTTTACTATCAGACAAGTCCACATTGGTACAGGGTCGTGTTCTTAGTAAACAACCAGTTT
TCGTACAATGTCTTAGAGCTGTGCCATCATGGTCCAACAACAGTGCCGTGGTGCATTTTA
GTAATGATGATTTTTGTCCTAATGTTACGGCAGAGGTGTTAAGGTTTAATCTGAATTTTA
GTGATACTGATGTTTATGTAGCTTCAAATAGTGATGATCAGTTGTACTTCACATTTGAAG
ACAATACTACTGCAGGCGTAGCTTGCTATAGTAGCGCTAATATTACAGATTATAAACCAA
ATACTAATGCTAGCTCACAGATCCCTTTTGGTAAAACCACGCATTCATACTTCTGTTTTG
CCAATTTCTCAAGCTCTCATGTTACACAATTTTTGGGTATATTGCCACCAACTGTGCGAG
AGTTTGCGTTTGGCAAAGACGGTTCCATTTTTGTTAATGGTTATAAATATTTTAGTTTTC
CACCTATCAAGAGTGTTAATTTCTCTATTAGTTCGGTTGAAAATTTTGGTTTTTGGACTA
TAGCTTACACCAATTACACCGATGTAATGGTGGATGTTAATGGCACTGGTATTACTAGGT
TGTTCTATTGTGACTCACCCATTAATAGAATTAAGTGCCAGCAACTGAAGTATGAACTAC
CAGATGGTTTTTATTCAGCTAGCATGCTTGTTAAAAAGGACTTACCTAAAACATTTGTAA
CTATGCCACAGTTTTATAATTGGATGAATGTCACTTTACATGTTGTGTTGAATGATACTG
AGAAAAAGGCTGATATTATTCTGGCTAAGGCTTCTGAATTAGCCTCACTTGCAGACGTTC
ACTTTGAAATTGCACAAGCTAATGGCAGTGTGGTTAACGCCACCAGTCTGTGTGTACAGA
CAAGACAGTTGGCTCTCTTCTATAAGTATACTAGCTTACAAGGTTTGTATACTTATTCTA

ATTTGGTAGAACTGCAGAATTATGACTGCCCATTTTCGCCACATCAATTTAATAATTATT
TGCAGTTTGAAACTTTGTGTTTTGATACAAGCCCGGCTGTTGCAGGTTGCAAGTGGTCTT
TAGTTCATGATGTTAAATGGCGCACACAGTTTGCAACAATTACAATCTCTTATAAAGATG
GTGCCAAGATTACAACTATGCCAAAAGCAAAGCTGGGTTTTCAAGATATCTCTAACATTG
TAAAAGATGAATGCACTGATTATAACATTTACGGATTTCAGGGCACTGGTATCATCAGAA
ACACCACCACTAGGATGGTTGCAGGCCTTTACTACACCTCAATTAGTGGTGACTTGCTTG
CATTTAAAAATAGCACTACAGGTGAAATTTTCACCGTAGTGCCATGTGGTTTAACGGCTC
AAGCTGCAGTGATTAATGATGAAATAGTTGGAGTTATAACAGCTGTTAATCAAACTGATC
TATTTGAGTTTGTAAATCATACACAAGCAAGAAGAGCACGCAGGTCAACAGGTTCACAGA
CCGTGCAAACTTATACTATGCCACAATTCTATTACATAACAAAGTGGAATAATGACAGCT
CTACTAATTGTACATCTGTCATTACGTACTCCTCTTTTGCTATTTGTAACACGGGTGAGA
TTAAATATGTTAATGTCACTCATGTGGAAATTGTAGATGATAGTATAGGTGTGATCAAAC
CTATCTCAACAGGCAACATATCCATACCTAAAAATTTCACTGTCGCAGTTCAAGCTGAAT
ACATTCAGATTCAAGTCAAACCCGTTGTGGTTGATTGCGCTAAGTATGTTTGTAATGGAA
ATAGTCATTGCCTCACACTACTTGCACAATACACATCTGCTTGTCAGACAATTGAAAATG
CCCTTAATTTTGGTGCACGTCTTGAATCGCTAATGCTTAATGATATGATCACAGTATCAG
ATCGTAGTTTGGAGCTTGCAACTGTTGAAAAATTTAACACCACGGTTTTAGGTGCTGAAA
AGTTGGGTGGTTTCTATTTTGATGGTTTGAGAGAGTTGTTACCACCTACAATTGGTAAGA
GGTCTGCTATTGAGGATTTACTGTTTAATAAGGTGGTAACTAGTGGCCTTGGAACTGTTG
ATGATGATTATAAAAAGTGCTCTGCTGGTACAGACGTTGCAGACCTTGTGTGTGCTCAAT
ATTATAATGGCATAATGGTACTACCTGGTGTTGTGGACCAAAATAAGATGTCTATGTACA
CCGCTTCCTTAATAGGTGGCATGGCCATGGGTTCTATCACATCAGCAGTAGCCGTCCCCT
TTGCTATGCAAGTGCAGGCCAGGCTTAATTATGTTGCATTACAAACGGATGTCTTACAGG
AAAACCAGAAAATACTTGCTAATGCCTTTAATAATGCCATCGGCAACATCACGCTAGCAC
TGGGAAAAGTTTCTAATGCCATTACTACTATTTCAGATGGATTTAATACAATGGCTTCAG
CACTTACTAAGATTCAGAGTGTGGTTAATCATCAGGGTGACGCATTGAGTCAACTTACTA
GTCAGTTACAGAAGAATTTCCAGGCTATTAGTAGTTCTATTGCTGAAATCTACAATAGAC
TAGAAAAGGTTGAGGCTGATGCTCAAGTTGACCGCCTCATTACTGGGCGATTGGCAGCAC
TTAATGCCTATGTGTCTCAAACTTTAACTCAATACTCAGAAGTCAAGGCTAGTAGGCAAT
TAGCAATGGAGAAGGTTAATGAGTGTGTTAAATCACAGTCTAATAGGTATGGGTTCTGTG
GAAATGGAACACACCTATTCTCTCTAGTCAATTCTGCACCTGATGGTTTACTCTTCTTTC
ACACAGTTTTGCTTCCTACAGAATGGGAAGTAGTGACGGCATGGTCAGGAATATGTGTTA
ATGACACGTATGCATATGTGTTGAAAGACTTTGAATATTCTATTTTCAGCTATAACGGTA
CGTATATGGTGACTCCTCGTAATATGTTTCAACCTAGAAAACCTCAGATGAGTGATTTCG
TGCAAATTACGAGTTGTGAAGTGACTTTTCTGAACACTACATATACGACATTTCAGGAGA
TTGTGATTGATTATATTGATATCAACAAGACTATCGCTGATATGCTTGAACAATATAATC
CCAATTACACAACACCTGAATTAAACTTACAGCTTGAAATCTTTAATCAGACAAAGTTAA
ACCTTACTGCAGAAATAGACCAATTAGAGCAAAGAGCAGACAACCTTACTAACATAGCAC
ATGAGCTACAACAGTACATTGACAATCTTAATAAGACCCTTGTTGACCTCGAATGGCTCA
ACAGAGTAGAGACTTATGTAAAATGGCCCTGGTATGTGTGGCTACTAATTGGATTAGTAG
TGGTCTTCTGCATACCATTGTTACTGTTTTGCTGTCTGAGTACTGGTTGTTGTGGATGCT
TTGGTTGTCTTGGCAGTTGTTGTCATTCTCTTTGCAGCAGACGACAATTTGAAAGTTATG
AACCCATTGAAAAGGTTCACATTCATTAAGAGACG

Unit 6

CGTCTCCATTAACTAAACGATTTATGGACACTGTCAAGTCCATTGGCATCTCTGTAGACG
CTGTACTTGACGAGTTAGATTCTATTGCTTTTGCTGTAACACTTAAAGTATTATTTACTT
CTGGTAGATTACTTGTGTGTATAGGTTTTGGTGATACTTTTGAGGAGGCTGAACAAAAAG
CTTATGCCAAACTTCAGCTGGATATTGAAGAGTCGTTTACTCATAAGACTGTTTAATATC
ACAGTTTATGATTTTTGTGTTAAAAACTGGTATAAGTTACCTTTTGCAGTTAGATTGCGT
ATCATAAATAATACAAAACCTAAAACAGCAAGTACTATAAAACGTAGAAGAAGTATTGTT
ATTGATTACAGAAAAATTGCCATTCTCAACGCGTCGCGAAAATGATTGGCGGACTGTTTC
TTAACACTCTAAGTTTTATAGTTACTACTCAACATGTGGTTGTTAATAACACACCACATT
TTAATACTGTAGTGCAACAACATCATGTTGTTAGTGCTAATGTTAAAAGTTTTCATCTAG
AGTTCAGCATAGCTGTTCTCTTTGTTTTATTTCTAGCTTTGTACCGTAGTACAAACTTTA

AAGCGTGTGTCGGTATCTTAATGTTTAAGATAGTATCAATGACACTTATAGGACCTATGC
TTATAGCATTTGGTCACTATATAGATGGCATTGTGACAACAACTGTCTTAGCTTTAAGAT
TCATTTACTTAGCATACTTTTGGTATGTTAATAGTAAATTTGAATTTATTTTATATAACA
CAACGACACTAATGTTTGTACATGACAGAGCTGCACCGTTTATGAGAAGTTCTCACGGCT
CTATTTATGTCACATTGTATGGAGGCATAAATTATATGTTCGTGAATGATCTTACGTTGC
ATTTTGTAGACCCTATGCTTGTAGGCATAGCTATACGTGGCCTAGTTCGTGCTGACCTAA
CAGTTGTTAGAGCAGTTGAACTTCTCAATGGTGATTTTATCTATATATTTTCACAGGAGG
CCGTCGTAGGAGTTTACAATGCAGCCTTTTCTCAGGCGGTTATAAACGAAATTGATTTAA
AAGAAGAAGAAGAACGTGTCTATGATGTTTCCTAGGGCATTTACTATCATAGACGACCAT
GGTATGGTTGTTAGTGTCTTCTTCTGGCTCCTGTTGATAATTGTATTGATATTGTTTTCA
ATAGCATTGCTAAATGTTATTAAATTGTGCATGGTATGTTGTAATTTGGGTAAGACTATT
ATAGTATTACCCGCTCGCCATGCATATGATGCCTATAAGACCTTTATGCAAATTAAGGCA
TATAATCCCGACGAAGCATTCTTAGTTTGAACTAAACAAAATGAAGTACATTTTGTTCTT
ACTTGCGTGCATAATTGCATGCGTTTATGGAGAACGTTATTGTGCCATGCAAAGTGCAAC
TTCTACAAGCTGCATCAATGGCACTGACCACAACTCATGTCAAACTTGCTTCGAACGTGG
TGATCTTATTTGGCATCTCGCTAACTGGAACTTCAGCTGGTCTGTAATATTGATTGTTTT
TATAACAGTGTTACAATACGGCAGACCGCAATTTAGCTGGCTCGTTTATGGCATTAAAAT
GCTGATCATGTGGCTATTATGGCCTATTGTTCTAGCGCTTACGATTTTTAATGCGTACTC
TGAGTACGAAGTTTCCAGATATGTAATGTTCGGCTTTAGTGTTGCAGGTGCAGTTGTAAC
GTTTGCATTATGGATGATGTATTTTGTGAGATCTATTCAGCTATATAGAAGGACCAAATC
ATGGTGGTCTTTTAACCCTGAGACTAATGCAATTCTCTGTGTTAATGCGTTGGGTAGAAG
CTACGTCCTACCCCTTGATGGCACTCCTACGGGTGTCACTCTCACGTTACTTTCAGGAAA
CCTATATGCTGAAGGGTTTAAAATGGCTGGTGGTCTTACCATCGAGCACTTGCCTAAATA
CGTCATGATTGCTACACCTAGTAGAACCATCGTTTATACCTTAGTTGGAAAACAGTTAAA
AGCAACTACTGCCACAGGATGGGCTTACTATGTAAAATCTAAAGCCGGTGATTACTCAAC
AGAAGCACGTACTGATAATTTGAGTGAACATGAAAAATTATTACATATGGTATAACTAAA
CTCCTAAATGGCCACACAGGGACAACGCGTCAACTGGGGAGATGAACCTTCCAAAAGACG
TGGTCGTTCTAACTCTCGTGGTCGGAAGAATAATGATATACCTTTATCATTCTTTAACCC
CATCACCCTCGAAAAAGGATCAAAATTTTGGAATTTATGTCCGAGAGACTTTGTTCCCAA
GGGAATAGGTAATAAAGATCAACAAATTGGTTATTGGAATAGACAAGCGCGTTACCGCAT
TGTCAAGGGTCAGCGTACTGATCTTCCAGAGAGATGGTTCTTCTATTTCTTAGGTACAGG
ACCTCACGCTGATGCTAAATTCAAAGACAAGATTGATGGAGTCTACTGGGTTGCTAGGGA
TGGAGCCACGAACAAGCCTACAACACTTGGCACTCGTGGAACCAATAATGAGTCCAAACC
ATTGAAGTTTGACGGTAAGATACCACCGCAATTTCAACTTGAAGTAAACCGTTCTAGGAA
CAATTCAAGGTCTGGTTCTCAGTCTAGATCTGTCTCAAGAAACAGGTCTCAATCCAGAGG
AAGACAACAATCCAATAACCAGAATACTAATGTTGAAGATACAATTGTTGCTGTGCTTCA
AAAATTAGGTGTTACTGACAAGCAAAGGTCACGTTCTAAATCCAGTGAACGTAGTGATTC
AAAACCTAGAGATACAACACCTAAAAATGCCAACAAACACACCTGGAAGAAAACTGCAGG
TAAGGGAGATGTGACAAACTTCTATGGTGCTAGAAGTGCTTCAGCTAACTTTGGTGATAG
TGATCTCGTTGCCAATGGTAACGCTGCCAAATGCTACCCTCAGATAGCTGAATGCGTTCC
ATCAGTGTCTAGCATGCTCTTCGGTAGTCAATGGTCTGCTGAAGAAGATGGAGATCAAGT
GAAAGTCACGCTCACTCACACCTACTACCTGCCAAAGGATGATGCCAAAACCAGCCAGTT
CCTAGAACAGATTGATGCCTACAAACGGCCTTCACAGGTGGCTAAAGATCAGAAGACAAG
GAAATCCCGACACAAGTCCGCTGATAAGAAGCCCGAGGAACTGTCTGTTACTCTTGTAGA
GGCATACACAGATGTGTTTGATGACACACAGGTTGAGATGATTGATGAGGTTACGAACTA
AACGCATGCTCGTTTTCCTCCATGCTGTACTTGTAACAGCTTTAATCTTACTACTAATTG
GTAGAATCCAATTACTAGAAAGGTTGTTACTCAGTCATCTGCTTAATCTCACAACAGTCA
GTAATGTTTTAGGTGTGCCTGACAGTAGTCTGCGTGTAAATTGCTTACAGCTTTTGAGAC
CAGACTGCCTTGATTTTAATATCTTACATAAAGTTTTAGCAGAAACCAGATTACTAGTAG
TAGTACTGCGAGTGATCTTTCTAGTTCTTCTAGGGTTTTCCTGCTATACATTGTTGGGTG
CATTATTTTAACATCATGATTGCTGTATTCCTTGTGTGTATTCTTTTGGCTAATGGAATT
AAAGCTATTGATGTGCAACACGACCTTCATGAACATCCAGTCCTAACTTGGGAATTATTG
CAACATTTCATAGGAAGTACTCTTTACATCACAACACATCAGGTCTTAGCATTACCGCTT
GGATCTCGTGTTGAATGTGAGAGTATCGAAGGTTTTAATTGCACATGGCCTGGCTTTCAA
AATCCTGCACATGATCACATTGATTTTTATTTTGATCTTTCTAATCCTTTCTATTCCTTT
GTAGATAATTTCTATATAGTAGGTGAAGGAAATCAAAGAATTAATCTCAGATTAGTTGGT
GCTGTGCCAAAACAAAAGAGATTAAATGTTGGTTGTCATACATCATTTGCTGTTGACCTT

CCTTTTGGAACTCAGATCTACCATGACAGGGATTTTCAACTACCTGTTGATGGTAGACAT
CTAGAGTGTGCTCACAGAGTTTACTTTGTGAAGTATTGTCCACAAAACCTGCATGGTTAT
TGCTTTAATGAGAAGTTGAAAGTCTACGACTTGAAGCAACTCAGAAGCAAAAAGGTCTTC
GACAAACTCAACCAACATCATAAAACTGAGTTGTAAGGCAACCCGATGTTTAAAACTGGT
CTTTCCGAGGAATTACTGGTCATCGCGCTGTCTACTCTTGTACAGAATGGTAAGCACGTG
TAATAGGAGGTACAAGCAACCCTATTGCATATTAGGAAGTTTAGATTTGATTTGGCAATG
CTAGATTTAGTAATTTAGAGAAGTTTAAAGATCCGCTATGACGAGCCAACAATGGAAGAG
CTAACGTCTGGATCTAGTAATTGTTTAAAATGTAAAATTGTTTGAAAATTTTCCTTTTGA
TAGTGATACACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGGTGGCCGGCATGGT
CCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATTCCGAGGGGACCGTCCCCTCGGTAAT
GGCGAATGGGACCCAGGGCCC

Unit 5T2

CGTCTCCCATGATTGTGCTCGTAACTTGCCTCTTGTTGTTATGTTCATACCACACAGTTT
TGAGTACAACAAATAATGAATGCATACAAGTTAACGTAACACAATTGGCTGGCAATGAAA
ACCTTATCAGAGATTTTCTGTTTAGTAATTTTAAAGAAGAAGGAAGTGTAGTTGTTGGTG
GTTATTACCCTACAGAGGTGTGGTACAACTGCTCTAGAACAGCTCGAACTACTGCCTTTC
AGTATTTTAATAATATACATGCCTTTTATTTTGTTATGGAAGCCATGGAAAATAGCACTG
GTAATGCACGTGGTAAACCATTATTATTTCATGTGCATGGTGAGCCTGTTAGTGTTATTA
TATCGGCTTATAGGGATGATGTGCAACAAAGGCCCCTTTTAAAACATGGGTTAGTGTGCA
TAACTAAAAATCGCCATATTAACTATGAACAATTCACCTCCAACCAGTGGAATTCCACAT
GTACGGGTGCTGACAGAAAAATTCCTTTCTCTGTCATACCCACGGACAATGGAACAAAAA
TCTATGGTCTTGAGTGGAATGATGACTTTGTTACAGCTTATATTAGTGGTCGTTCTTATC
ACTTGAACATCAATACTAATTGGTTTAACAATGTCACACTTTTGTATTCACGCTCAAGCA
CTGCTACCTGGGAATACAGTGCTGCATATGCTTACCAAGGTGTTTCTAACTTCACTTATT
ACAAGTTAAATAACACCAATGGTCTAAAAACCTATGAATTATGTGAAGATTATGAACATT
GCACTGGCTATGCTACCAATGTATTTGCTCCGACATCAGGTGGTTACATACCTGATGGAT
TTAGTTTTAACAATTGGTTCTTGCTTACAAATAGTTCCACTTTTGTTAGTGGCAGGTTTG
TAACAAATCAACCATTATTGATTAATTGCTTGTGGCCAGTGCCCAGTTTTGGTGTAGCAG
CACAAGAATTTTGTTTTGAAGGTGCACAGTTTAGCCAATGTAATGGTGTGTCTTTAAATA
ACACAGTGGATGTTATTAGATTCAACCTTAATTTCACTGCAGATGTACAATCTGGTATGG
GTGCTACAGTATTTTCACTGAATACAACAGGTGGTGTCATTCTTGAAATTTCATGTTATA
GTGACACAGTGAGTGAGTCTAGTTCTTACAGTTATGGTGAAATCCCGTTCGGCATAACTG
ACGGACCACGATACTGTTATGTACTTTACAATGGCACAGCTCTTAAATATTTAGGAACAT
TACCACCCAGTGTAAAGGAAATTGCTATTAGTAAGTGGGGCCATTTTTATATTAATGGTT
ACAATTTCTTTAGCACATTTCCTATTGGTTGTATATCTTTTAATTTAACCACTGGTGTTA
GTGGAGCTTTTTGGACAATTGCTTACACATCGTATACTGAAGCATTAGTACAAGTTGAAA
ACACAGCTATTAAAAATGTGACGTATTGTAACAGTCACATTAATAACATTAAATGTTCTC
AACTTACTGCTAATTTGAATAATGGATTTTATCCTGTTGCTTCAAGTGAAGTAGGTTTCG
TTAATAAGAGTGTTGTGTTATTACCTAGCTTTTTCACATACACCGCTGTCAATATAACCA
TTGATCTTGGTATGAAGCTTAGTGGTTATGGTCAACCCATAGCCTCGACACTAAGTAACA
TCACACTACCAATGCAGGATAACAATACTGATGTGTACTGTATTCGTTCTAACCAATTCT
CAGTTTATGTTCATTCCACTTGCAAAAGTTCTTTATGGGACAATATTTTTAATCAAGACT
GCACGGATGTTTTAGAGGCTACGGCTGTTATAAAAACTGGTACTTGTCCTTTCTCATTTG
ATAAATTGAACAATTACTTGACTTTTAACAAGTTCTGTTTGTCGTTGAGTCCTGTTGGTG
CTAATTGCAAGTTTGATGTTGCTGCACGTACAAGAACCAATGAGCAGGTTGTTAGAAGTC
TATATGTAATATATGAAGAAGGAGACAACATAGTGGGTGTACCGTCTGATAATAGCGGTC
TGCACGATTTGTCTGTGCTACACCTAGACTCCTGTACAGATTACAATATATATGGTAGAA
CTGGTGTTGGTATTATTAGACGAACTAACAGTACGCTACTTAGTGGCTTATATTACACAT
CACTATCAGGTGATTTGTTAGGCTTTAAAAATGTTAGTGATGGTGTCATTTATTCTGTGA
CGCCATGTGATGTAAGCGTACAAGCGGCTGTTATTGATGGTGCCATAGTTGGAGCTATGA
CTTCCATTAACAGTGAACTGTTAGGTCTAACACATTGGACAACGACACCTAATTTTTATT
ACTACTCTATATATAATTATACAAGTGAGAGGACTCGTGGCACTGCAATTGACAGTAACG
ATGTTGATTGTGAACCTGTCATAACCTATTCTAATATAGGTGTTTGTAAAAATGGTGCTT
TGGTTTTTATTAACGTCACACATTCTGACGGtGAtGTGCAACCAATTAGCACTGGTAATG

TCACGATACCTACAAATTTTACTATATCTGTGCAAGTTGAATACATGCAGGTTTACACTA
CACCAGTATCAATAGATTGTGCAAGATACGTTTGTAATGGTAACCCTAGATGTAACAAAT
TGTTAACACAATATGTGTCTGCATGTCAAACTATTGAACAAGCACTTGCAATGGGTGCCA
GACTTGAAAACATGGAGGTTGATTCCATGTTGTTTGTCTCGGAAAATGCCCTTAAATTGG
CATCTGTTGAGGCGTTCAATAGTACAGAAAATTTAGATCCTATTTACAAAGAATGGCCTA
GCATAGGTGGTTCTTGGCTAGGAGGTCTAAAAGATATACTACCGTCCCATAATAGCAAAC
GTAAGTATGGTTCTGCTATAGAAGATTTGCTTTTTGATAAAGTTGTAACATCTGGTTTAG
GTACAGTTGATGAAGATTATAAACGTTGTACTGGTGGTTACGACATAGCAGACTTGGTGT
GTGCTCAATATTACAATGGCATCATGGTTCTACCAGGTGTAGCTAATGCTGACAAGATGA
CTATGTACACAGCATCACTTGCAGGTGGTATAACATTAGGTGCACTTGGTGGTGGCGCCG
TGGCTATACCTTTTGCAGTAGCAGTACAGGCTAGACTTAATTATGTTGCTCTACAAACTG
ATGTATTGAATAAAAACCAACAGATCCTGGCTAATGCTTTCAATCAAGCTATTGGTAACA
TTACACAGGCTTTTGGTAAGGTTAATGATGCTATACATCAAACATCACAAGGTCTTGCCA
CTGTTGCTAAAGCGTTGGCAAAAGTGCAAGATGTTGTCAACACACAAGGGCAAGCTTTAA
GTCACCTTACAGTACAATTGCAAAATAATTTTCAAGCCATTAGTAGTTCTATTAGTGATA
TTTATAACAGGCTTGACGAACTGAGTGCTGATGCACAAGTTGATAGGCTGATTACAGGTA
GACTTACAGCACTTAATGCATTTGTGTCTCAGACTCTAACCAGACAAGCAGAGGTTAGGG
CTAGTAGACAACTTGCCAAAGACAAGGTTAATGAATGTGTTAGGTCTCAGTCTCAGAGAT
TCGGATTCTGTGGTAATGGTACACATTTGTTTTCACTAGCAAATGCAGCACCAAATGGCA
TGATTTTCTTTCATACAGTACTATTACCAACAGCTTATGAAACTGTAACAGCTTGGTCAG
GTATTTGTGCTTCAGATGGCGATCGCACTTTCGGACTTGTCGTTAAAGATGTGCAGTTGA
CGTTGTTTCGTAATCTAGATGACAAGTTCTATTTGACCCCCAGAACTATGTATCAGCCTA
GAGTTGCAACTAGTTCTGATTTTGTTCAAATTGAAGGGTGTGATGTGTTGTTTGTCAACG
CGACTGTAATTGATTTGCCTAGTATTATACCTGACTATATTGACATTAATCAAACTGTTC
AAGACATATTAGAAAATTACGAACCAAACTGGACTGTACCTGAATTTACACTTGATATTT
TCAACGCAACCTATTTAAATCTGACTGGTGAAATTGATGACTTAGAGTTTAGGTCAGAAA
AGCTACATAACACTACAGTAGAACTTGCCATTCTCATTGATAACATTAATAATACATTAG
TCAATCTTGAATGGCTCAATAGAATTGAAACTTATGTAAAATGGCCTTGGTATGTGTGGC
TACTGATAGGTTTAGTAGTAGTATTTTGCATACCATTACTGCTATTTTGCTGTTTTAGCA
CAGGTTGTTGTGGATGCATAGGTTGTTTAGGAAGTTGTTGTCACTCTATATGTAGTAGAA
GACAATTTGAAAATTATGAACCAATTGAAAAAGTGCATGTCCATTAAGAGACG

Unit 5R

CGTCTCCCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCG
AGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATG
CCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCT
GGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACC
ACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCA
CCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCG
ACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCC
TGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGC
AGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGC
AGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG
ACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGTCCGGACTCAGATCTAGGcGtCGACCTTCCATGACCGAGTACAAGCCCACGGTGC
GCCTCGCCACCCGCGACGACGTCCCCAGGGCCGTACGCACCCTCGCCGCCGCGTTCGCCG
ACTACCCCGCCACGCGCCACACCGTCGATCCGGACCGCCACATCGAGCGGGTCACCGAGC
TGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGGAAGGTGTGGGTCGCGGACG
ACGGCCCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCG
CCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGA
TGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCG
GtGTaTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGG
AGGCGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCC
CCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGC

GCACCTGGTGCATGACCCGCAAGCCCGGTGCCTAAATTAAGAGACG

Unit 6R

CGTCTCCATTAACTAAACGATTTATGGACACTGTCAAGTCCATTGGCATCTCTGTAGACG
CTGTACTTGACGAGTTAGATTCTATTGCTTTTGCTGTAACACTTAAAGTATTATTTACTT
CTGGTAGATTACTTGTGTGTATAGGTTTTGGTGATACTTTTGAGGAGGCTGAACAAAAAG
CTTATGCCAAACTTCAGCTGGATATTGAAGAGTCGTTTACTCATAAGACTGTTTAATATC
ACAGTTTATGATTTTTGTGTTAAAAACTGGTATAAGTTACCTTTTGCAGTTAGATTGCGT
ATCATAAATAATACAAAACCTAAAACAGCAAGTACTATAAAACGTAGAAGAAGTATTGTT
ATTGATTACAGAAAAATTGCCATTCTCAACGCGTCGCGAAAATGATTGGCGGACTGTTTC
TTAACACTCTAAGTTTTATAGTTACTACTCAACATGTGGTTGTTAATAACACACCACATT
TTAATACTGTAGTGCAACAACATCATGTTGTTAGTGCTAATGTTAAAAGTTTTCATCTAG
AGTTCAGCATAGCTGTTCTCTTTGTTTTATTTCTAGCTTTGTACCGTAGTACAAACTTTA
AAGCGTGTGTCGGTATCTTAATGTTTAAGATAGTATCAATGACACTTATAGGACCTATGC
TTATAGCATTTGGTCACTATATAGATGGCATTGTGACAACAACTGTCTTAGCTTTAAGAT
TCATTTACTTAGCATACTTTTGGTATGTTAATAGTAAATTTGAATTTATTTTATATAACA
CAACGACACTAATGTTTGTACATGACAGAGCTGCACCGTTTATGAGAAGTTCTCACGGCT
CTATTTATGTCACATTGTATGGAGGCATAAATTATATGTTCGTGAATGATCTTACGTTGC
ATTTTGTAGACCCTATGCTTGTAGGCATAGCTATACGTGGCCTAGTTCGTGCTGACCTAA
CAGTTGTTAGAGCAGTTGAACTTCTCAATGGTGATTTTATCTATATATTTTCACAGGAGG
CCGTCGTAGGAGTTTACAATGCAGCCTTTTCTCAGGCGGTTATAAACGAAATTGATTTAA
AAGAAGAAGAAGAACGTGTCTATGATGTTTCCTAGGGCATTTACTATCATAGACGACCAT
GGTATGGTTGTTAGTGTCTTCTTCTGGCTCCTGTTGATAATTGTATTGATATTGTTTTCA
ATAGCATTGCTAAATGTTATTAAATTGTGCATGGTATGTTGTAATTTGGGTAAGACTATT
ATAGTATTACCCGCTCGCCATGCATATGATGCCTATAAGACCTTTATGCAAATTAAGGCA
TATAATCCCGACGAAGCATTCTTAGTTTGAACTAAACAAAATGACTTCGAAAGTTTATGA
TCCAGAACAAAGGAAACGGATGATAACTGGTCCGCAGTGGTGGGCCAGATGTAAACAAAT
GAATGTTCTTGATTCATTTATTAATTATTATGATTCAGAAAAACATGCAGAAAATGCTGT
TATTTTTTTACATGGTAACGCGGCCTCTTCTTATTTATGGCGACATGTTGTGCCACATAT
TGAGCCAGTAGCGCGGTGTATTATACCAGACCTTATTGGTATGGGCAAATCAGGCAAATC
TGGTAATGGTTCTTATAGGTTACTTGATCATTACAAATATCTTACTGCATGGTTTGAACT
TCTTAATTTACCAAAGAAGATCATTTTTGTCGGCCATGATTGGGGTGCTTGTTTGGCATT
TCATTATAGCTATGAGCATCAAGATAAGATCAAAGCAATAGTTCACGCTGAAAGTGTAGT
AGATGTGATTGAATCATGGGATGAATGGCCTGATATTGAAGAAGATATTGCGTTGATCAA
ATCTGAAGAAGGAGAAAAAATGGTTTTGGAGAATAACTTCTTCGTGGAAACCATGTTGCC
ATCAAAAATCATGAGAAAGTTAGAACCAGAAGAATTTGCAGCATATCTTGAACCATTCAA
AGAGAAAGGTGAAGTTCGTCGTCCAACATTATCATGGCCTCGTGAAATCCCGTTAGTAAA
AGGTGGTAAACCTGACGTTGTACAAATTGTTAGGAATTATAATGCTTATCTACGTGCAAG
TGATGATTTACCAAAAATGTTTATTGAATCGGACCCAGGATTCTTTTCCAATGCTATTGT
TGAAGGTGCCAAGAAGTTTCCTAATACTGAATTTGTCAAAGTAAAAGGTCTTCATTTTTC
GCAAGAAGATGCACCTGATGAAATGGGAAAATATATCAAATCGTTCGTTGAGCGAGTTCT
CAAAAATGAACAATAATATAACTAAACTCCTAAATGGCCACACAGGGACAACGCGTCAAC
TGGGGAGATGAACCTTCCAAAAGACGTGGTCGTTCTAACTCTCGTGGTCGGAAGAATAAT
GATATACCTTTATCATTCTTTAACCCCATCACCCTCGAAAAAGGATCAAAATTTTGGAAT
TTATGTCCGAGAGACTTTGTTCCCAAGGGAATAGGTAATAAAGATCAACAAATTGGTTAT
TGGAATAGACAAGCGCGTTACCGCATTGTCAAGGGTCAGCGTACTGATCTTCCAGAGAGA
TGGTTCTTCTATTTCTTAGGTACAGGACCTCACGCTGATGCTAAATTCAAAGACAAGATT
GATGGAGTCTACTGGGTTGCTAGGGATGGAGCCACGAACAAGCCTACAACACTTGGCACT
CGTGGAACCAATAATGAGTCCAAACCATTGAAGTTTGACGGTAAGATACCACCGCAATTT
CAACTTGAAGTAAACCGTTCTAGGAACAATTCAAGGTCTGGTTCTCAGTCTAGATCTGTC
TCAAGAAACAGGTCTCAATCCAGAGGAAGACAACAATCCAATAACCAGAATACTAATGTT
GAAGATACAATTGTTGCTGTGCTTCAAAAATTAGGTGTTACTGACAAGCAAAGGTCACGT
TCTAAATCCAGTGAACGTAGTGATTCAAAACCTAGAGATACAACACCTAAAAATGCCAAC
AAACACACCTGGAAGAAAACTGCAGGTAAGGGAGATGTGACAAACTTCTATGGTGCTAGA
AGTGCTTCAGCTAACTTTGGTGATAGTGATCTCGTTGCCAATGGTAACGCTGCCAAATGC

TACCCTCAGATAGCTGAATGCGTTCCATCAGTGTCTAGCATGCTCTTCGGTAGTCAATGG
TCTGCTGAAGAAGATGGAGATCAAGTGAAAGTCACGCTCACTCACACCTACTACCTGCCA
AAGGATGATGCCAAAACCAGCCAGTTCCTAGAACAGATTGATGCCTACAAACGGCCTTCA
CAGGTGGCTAAAGATCAGAAGACAAGGAAATCCCGACACAAGTCCGCTGATAAGAAGCCC
GAGGAACTGTCTGTTACTCTTGTAGAGGCATACACAGATGTGTTTGATGACACACAGGTT
GAGATGATTGATGAGGTTACGAACTAAACGCATGCTCGTTTTCCTCCATGCTGTACTTGT
AACAGCTTTAATCTTACTACTAATTGGTAGAATCCAATTACTAGAAAGGTTGTTACTCAG
TCATCTGCTTAATCTCACAACAGTCAGTAATGTTTTAGGTGTGCCTGACAGTAGTCTGCG
TGTAAATTGCTTACAGCTTTTGAGACCAGACTGCCTTGATTTTAATATCTTACATAAAGT
TTTAGCAGAAACCAGATTACTAGTAGTAGTACTGCGAGTGATCTTTCTAGTTCTTCTAGG
GTTTTCCTGCTATACATTGTTGGGTGCATTATTTTAACATCATGATTGCTGTATTCCTTG
TGTGTATTCTTTTGGCTAATGGAATTAAAGCTATTGATGTGCAACACGACCTTCATGAAC
ATCCAGTCCTAACTTGGGAATTATTGCAACATTTCATAGGAAGTACTCTTTACATCACAA
CACATCAGGTCTTAGCATTACCGCTTGGATCTCGTGTTGAATGTGAGAGTATCGAAGGTT
TTAATTGCACATGGCCTGGCTTTCAAAATCCTGCACATGATCACATTGATTTTTATTTTG
ATCTTTCTAATCCTTTCTATTCCTTTGTAGATAATTTCTATATAGTAGGTGAAGGAAATC
AAAGAATTAATCTCAGATTAGTTGGTGCTGTGCCAAAACAAAAGAGATTAAATGTTGGTT
GTCATACATCATTTGCTGTTGACCTTCCTTTTGGAACTCAGATCTACCATGACAGGGATT
TTCAACTACCTGTTGATGGTAGACATCTAGAGTGTGCTCACAGAGTTTACTTTGTGAAGT
ATTGTCCACAAAACCTGCATGGTTATTGCTTTAATGAGAAGTTGAAAGTCTACGACTTGA
AGCAACTCAGAAGCAAAAAGGTCTTCGACAAACTCAACCAACATCATAAAACTGAGTTGT
AAGGCAACCCGATGTTTAAAACTGGTCTTTCCGAGGAATTACTGGTCATCGCGCTGTCTA
CTCTTGTACAGAATGGTAAGCACGTGTAATAGGAGGTACAAGCAACCCTATTGCATATTA
GGAAGTTTAGATTTGATTTGGCAATGCTAGATTTAGTAATTTAGAGAAGTTTAAAGATCC
GCTATGACGAGCCAACAATGGAAGAGCTAACGTCTGGATCTAGTAATTGTTTAAAATGTA
AAATTGTTTGAAAATTTTCCTTTTGATAGTGATACACAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAACCGGTGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATTC
CGAGGGGACCGTCCCCTCGGTAATGGCGAATGGGACCCA

pHSPA1A sequence

NNNN Feline HSPA1A coding sequence

NNNN 3x FLAG Tag

NNNN Restriction site

CAAGCGGCCGCACCATGGCT
AAAACCGCAGCCATCGGCATCGACCTGGGCACCACCTACTCATGCGTGGGGGTGTTCCAGCACGGCAAGG
TAGAGATCATCGCCAACGACCAGGGCAACCGCACCACCCCCAGCTACGTGGCCTTCACGGACACCGAGCG
GCTCATCGGGGATGCGGCCAAGAACCAGGTGGCGCTGAACCCGCAGAACACCGTGTTCGACGCGAAGCGG
CTGATCGGCCGCAAGTTCGGCGACCCGGTGGTGCAGTCGGATATGAAGCACTGGCCTTTCCGGGTGATCA
ACGACGGAGATAAGCCCAAGGTGCAGGTGAGCTACAAGGGGGAGACCAAGGCGTTCTACCCCGAGGAGAT
CTCGTCCATGGTGCTGACCAAGATGAAGGAGATCGCCGAGGCGTACCTGGGCTACCCGGTGACCAACGCG
GTGATCACCGTGCCGGCCTACTTCAACGACTCGCAGCGGCAGGCCACCAAGGACGCGGGGGTGATCGCGG
GGCTGAACGTGCTGCGGATCATCAACGAGCCCACGGCCGCCGCCATCGCCTACGGCCTGGACAGGACGGG
CAAGGGGGAGCGCAACGTGCTCATCTTTGACCTGGGCGGGGGCACCTTCGACGTGTCCATCCTGACGATC
GACGACGGCATCTTCGAGGTGAAGGCCACGGCGGGGGACACCCACCTGGGTGGGGAGGACTTTGACAACA
GGCTGGTGAACCACTTCGTGGAGGAGTTCAAGAGGAAGCACAAGAAGGACATCAGCCAGAACAAGCGAGC
CGTGAGGCGGCTGCGCACCGCCTGCGAGCGGGCCAAGAGGACCCTGTCGTCCAGCACCCAGGCCAGCCTG
GAGATCGACTCCCTGTTCGAGGGCATCGACTTCTACACGTCCATCACGAGGGCGCGGTTCGAGGAGCTGT
GCTCGGACCTGTTCCGGAGCACCCTGGAGCCGGTGGAGAAGGCTCTGCGCGACGCCAAGCTGGACAAGGC
CCAGATCCACGACCTGGTCCTGGTGGGGGGCTCCACCCGCATCCCCAAGGTGCAGAAACTGCTGCAAGAC
TTCTTCAACGGGCGCGATCTCAACAAGAGCATCAACCCCGACGAGGCAGTGGCCTACGGGGCGGCAGTGC
AGGCGGCCATCCTGATGGGGGACAAGTCTGAGAACGTGCAGGACCTGTTGCTGCTGGACGTGGCGCCCCT
GTCGCTGGGGCTGGAGACGGCCGGCGGCGTGATGACTGCCCTGATCAAGCGCAACTCCACCATCCCCACC
AAGCAGACGCAGATCTTCACCACCTACTCGGACAACCAGCCCGGAGTGCTGATCCAGGTGTACGAGGGTG
AGAGGGCCATGACGCGGGACAATAACCTGCTGGGGCGCTTCGAGCTGAGCGGCATCCCCCCGGCCCCACG
GGGAGTGCCCCAGATCGAGGTGACCTTCGACATCGATGCCAATGGCATCCTGAACGTCACGGCCACGGAC
AAGAGCACGGGCAAAGCCAACAAGATCACCATCACCAACGACAAGGGCCGGCTGAGCAAGGAGGAGATCG
AGCGCATGGTGCAGGAGGCGGAGAAGTACAAAGCCGAGGACGAGGTGCAGCGCGAGAGGGTGTCTGCCAA
GAACGCGCTGGAGTCGTACGCCTTCAACATGAAGAGTGCCGTGGAGGATGAGGGTCTCAAGGGGAAGATC
AGCGAGGCCGACAAGAAGAAGGTGCTGGACAAGTGCCAGGAGGTCATTTCCTGGCTGGACGCCAACACCT
TGGCCGAGAAGGACGAGTTTGAGCACAAGAGGAAGGAGCTGGAGCAGGTGTGTAACCCCATCATCAGCGG
ACTGTACCAGGGGGCGGGTGGCCCTGGGGCTGGTGGCTTTGGGGCTCAGGCCCCCAGAGGGGGCTCTGGG
TCAGGCCCCACCATTGAGGAGGTGGATGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCG
ATTACAAGGATGACGATGACAAG

Primers and oligonucleotides

Primer name	Sequence	Use
U1-1	AAATCGCCCGTACTGGAAATG	Sequencing
U1-2	TTGAAGGCGAATTGAACGAC	Sequencing
U1-3	ATTGTCATTGTAGGCGACATGG	Sequencing
U1-4	GTGACACAGCTCGTAAAGAC	Sequencing
U1-5	AAAGTGATTCTGATGCAAGCG	Sequencing
U1-6	AAGCGTGGAAGGTTCAGGTG	Sequencing
U1-7	TTCACAGGTGGTAAGTTGACG	Sequencing
U2-1	GAGTGGTGCAGATTACACAC	Sequencing
U2-1 Rev	CTGCCTAGCAATCTTAGAACAAG	Sequencing
U2-2	TTAAGGCAGTTCTCGCACTTAAAC	Sequencing
U2-3	TGGCTGTTAATAACGCTCATAG	Sequencing
U2-4	TGGTAGATCCCTAGTGTTTGCG	Sequencing
U2-5	TTGGTGATTGTACTCTCCTAGTTG	Sequencing
U2-6	CCTAGACACGTCATTGCAAG	Sequencing
U2-7	TTTGGAGGACGCACTATACTC	Sequencing
U3-1	ATTGTGTTGGGCTGCATAATG	Sequencing
U3-2	TTGAACCTGACCTGGCCTTTG	Sequencing
U3-3	TTGTTCAAGTACCAACCGGC	Sequencing
U3-4	CGTGGAGAAAGGTTACATAGG	Sequencing
U3-5	CTTTGAGGAGGGATCTGAGTTG	Sequencing
U3-6	GCCTAATATGATACGAATGGCTTC	Sequencing
U3-7	ACCTTAATGTTGGACCACACG	Sequencing
U4-1	AACTATCCAAGGTCCTCCTG	Sequencing
U4-1 Rev	ACAGAGAATACCAACCTTTGCTC	Sequencing
U4-2	ATCTCACCCTATAACAGTCAGAAC	Sequencing
U4-3	ACCACCGGGTGAGCAATTTG	Sequencing
U4-4	GATGTCTGGAGTATGATTACATGG	Sequencing
U4-5	CAAACTTGGACTCACACCTC	Sequencing
U4-6	GCATCTTCTTATATCACAAGTGCG	Sequencing
U4-7	ATGATATGAGGGATTACGTTTCCG	Sequencing
U5-1	CCAGTATTTCCTACCACTGGAAC	Sequencing
U5-2	ACTGTGCGAGAGTTTGCGTTTGGC	Sequencing
U5-3	TGTTAAATGGCGCACACAGTTTGC	Sequencing
U5-4	ACATTCAGATTCAAGTCAAACCCG	Sequencing
U5-5	ATGCCTTTAATAATGCCATCGG	Sequencing
U5-6	TTCCTACAGAATGGGAAGTAGTG	Sequencing
U6-1	TTGTGACAACAACTGTCTTAGC	Sequencing
U6-2	ATTAAGGCATATAATCCCGACG	Sequencing
U6-3	GTGGTCTTACCATCGAGCAC	Sequencing
U6-4	GGAACAATTCAAGGTCTGGTTCTC	Sequencing
U6-5	ATGACACACAGGTTGAGATGATTG	Sequencing
U6-6	AATCTCAGATTAGTTGGTGCTGTG	Sequencing
U6-6 II	ACATCATTTGCTGTTGACCTTCC	Sequencing

U5R-1	GCATGGACGAGCTGTACAAG	Sequencing/Amp 6.2 PCR
U6R-1	TTGTGACAACAACTGTCTTAGC	Sequencing
U6R-2	AGGAAACGGATGATAACTGGTC	Sequencing
U6R-3	TACGTGCAAGTGATGATTTACC	Sequencing
U6R-3 Rev	TGGATTTAGAACGTGACCTTTG	Sequencing
U5T2-1	ACCAAGGTGTTTCTAACTTCAC	Sequencing
U5T2-2	TTGGACAATTGCTTACACATCG	Sequencing
U5T2-3	GTCTGTGCTACACCTAGACTCC	Sequencing
U5T2-4	CTTAAATTGGCATCTGTTGAGGC	Sequencing
U5T2-5	ATGCACAAGTTGATAGGCTG	Sequencing
U5T2-6	GTGCAGTTGACGTTGTTTCG	Sequencing
Twist For	CCGCACGCATCTGGAATAAGG	Sequencing/PCR (Twist vector)
Twist Rev	CATGCATCCACCATCGCAGAC	Sequencing/PCR (Twist vector)
M13-UC For	CCCAGTCACGACGTTGTAAAACG	Sequencing/PCR (pWSK29 vector)
M13-UC Rev	AGCGGATAACAATTTCACACAGG	Sequencing/PCR (pWSK29 vector)
Unit 4 For	TAAGCAGCGGCCGCCGTCTCCTGATGTAACAAAATTG	Unit 4 PCR
Unit 4 Rev	TGCTTAGTCGACCGTCTCTCATGGTGTGTTAAC	Unit 4 PCR
N gene For	TAAGCAGTCGACTAATACGACTCACTATAGATGGCCACACAGGGACAAC	N gene PCR
N gene Rev	TGCTTATCTAGATTTTTTTTTTTTTTTTTTTTTTTTTTAGTTCGTAACCTCATCAATC	N gene PCR
FL For	GGCCCTAATACGACTCACTATAGAC	Long range PCR
FL Rev	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGTATCAC	Long range PCR
F1	AAAGTTAAGTGAGTGTAGCGTGG	Amp 1 PCR
R1	GTGCGAGAACTGCCTTAA	Amp 1 PCR
F2	GATTGAACTCACGTGGTCATT	Amp 2 PCR
R2	GAGGTCTTCATCTGAACCCAC	Amp 2 PCR
F3	GCTAGTGTTAGAAATGTCTGTGTT	Amp 3 PCR
R3	AAAAGCTCTACTAACGTGGTC	Amp 3 PCR
F4	CATCCTGCAATTGATGGATTG	Amp 4 PCR
R4	TCCGGGTACATGTCTACGTTG	Amp 4 PCR
F5	GATTGGTCCATTGTGTACCC	Amp 5 PCR
R5	CACTTGTACAAAACACAGTCC	Amp 5 PCR
F6	GTATTAAGAAGATGGTTGCCAG	Amp 6/6.1 PCR
R6	ATAACCGCCTGAGAAAAGGCT	Amp 6/6.2 PCR
R6Rep	TGAGGAAGAGTTCTTGCAGC	Amp 6.1 PCR
F7	CATTGGCATCTCTGTAGACG	Amp 7 PCR
R7	CACTATCAAAAGGAAAATTTTC	Amp 7 PCR

Oligonucleotide adaptors for pWSK29 vector

U1_F	TCGACACTTGGGCCCTAATACGACTCACTAGGATCCTTTAGCATTACAAAGGCTTGAGACGGC
U1_R	GGCCGCCGTCTCAAGCCTTTGTAATGCTAAAGGATCCTAGTGAGTCGTATTAGGGCCCAAGTG
U5_F	TCGACCGTCTCCCATGATACTATTGTTGGTGGATCCAAAGGTTCACATTCATTAAGAGACGGC
U5_R	GGCCGCCGTCTCTTAATGAATGTGAACCTTTGGATCCACCAACAATAGTATCATGGGAGACGG
U5T2_F	TCGACCGTCTCCCATGATTGTGCTCGTAACGGATCCAAAAGTGCATGTCCATTAAGAGACGGC
U5T2_R	GGCCGCCGTCTCTTAATGGACATGCACTTTTGGATCCGTTACGAGCACAATCATGGGAGACGG

Synthetic DNA fragment to repair Unit 4

ACTAGTTTGCCTACGAATGTGGCTTTTGAGCTGTATGCAAAACGCAAACTTGGACTCACACCTCCGTTAACAATA
CTTAGAAATCTAGGTGTTGTCGCAACACATAAATTCGTGTTGTGGGATTATGAAGCCGAATGTCCTTTCTCAAAC
TTTACTAAGCAAGTGTGTGCGTACACTGATCTTGATGGTGAAGTTGTAACATGTTTTGATAATAGTATTAGTGGT
TCTTTTGAACGCTTTACTACTACGAAAGATGCAGTGCTTATTTCTAATAACGCTGTGAGGGGACTCAGTGCCATT
AAATTACAATATGGCTTTTTGAATGATTTACCTGTAAGTACTGTAGGAAACAAACCTGTCACATGGTACATCTAT
GTGCGCAAGAACGGTGAGTACGTCGAACAGATTGACAGTTATTACACACAGGGACGTACTTTTGAAACCTTTAAA
CCTCGTAGTACAATGGAAGAAGACTTTTTTAGTATGGATACTACACTCTTCATCCAAAAGTATGGTCTTGAGGAT
TATGGTTTTGAACATGTTGTATTCGGAGATGTTTCTAAAACTACCATTGGTGGTATGCATCTTCTTATATCACAA
GTGCGCCTAGCAAAAATGGGTTTGTTTTCTGTCCAAGAATTTATGAGTAATTCTGACAGTACACTGAAAAGTTGT
TGTATAACATACGCTGACGATCCGTCTTCTAAGAATGTGTGCACTTATATGGACATACTCTTAGACGATTTTGTG
ACTATCGTTAAGGGCTTAGATCT

Appendix B: Thermocycler programmes

OneTaq DNA polymerase PCR

Step	Temperature	Time
Denaturation	94 °C	30 seconds
30 cycles	94 °C	30 seconds
	45-68 °C	1 minute
	68 °C	1 minute per kb
Final extension	68 °C	5 minutes
Hold	4 °C	Infinite

Phusion DNA polymerase PCR (two step)

Step	Temperature	Time
Denaturation	98 °C	30 seconds
30 cycles	98 °C	10 seconds
	72 °C	30 seconds per kb
Final extension	72 °C	10 minutes
Hold	4 °C	Infinite

Phusion DNA polymerase PCR (three step)

Step	Temperature	Time
Denaturation	98 °C	30 seconds
30 cycles	98 °C	10 seconds
	45-72 °C	30 seconds
	72 °C	30 seconds per kb
Final extension	72 °C	10 minutes
Hold	4 °C	Infinite

GoTaq DNA polymerase long range PCR

Step	Temperature	Time
Denaturation	95 °C	2 minutes
30 cycles	93 °C	20 seconds
	52 °C	30 seconds
	65 °C	1 minute per kb
Final extension	72 °C	10 minutes
Hold	4 °C	Infinite

SuperScript IV One-step RT-PCR

Step	Temperature	Time
Reverse transcription	50 °C	10 minutes
Denaturation	98 °C	2 minutes
35 cycles	98 °C	10 seconds
	55-72 °C	10 seconds
	72 °C	30 seconds per kb
Final extension	72 °C	5 minutes
Hold	4 °C	Infinite

Peptide number	Sequence
19	H-TSAVAVPFAMQVQAR-OH
20	H-VPFAMQVQARLNYVA-OH
21	H-QVQARLNYVALQTDV-OH
22	H-LNYVALQTDVLQENQ-OH
23	H-LQTDVLQENQKILAN-OH
24	H-LQENQKILANAFNNA-OH
25	H-KILANAFNNAIGNIT-OH
26	H-AFNNAIGNITLALGK-OH
27	H-IGNITLALGKVSNAI-OH
28	H-LALGKVSNAITTISD-OH
29	H-LALGKVSNAITTTSD-OH
30	H-VSNAITTISDGFNTM-OH
31	H-VSNAITTTSDGFNTM-OH
32	H-AITTTSDGFNTMTTI-OH
33	H-AITTISDGFNTMTTI-OH
34	H-TTISDGFNTMASALT-OH
35	H-GFNTMASALTKIQSV-OH
36	H-ASALTKIQSVVNHQG-OH
37	H-KIQSVVNHQGDALSQ-OH
38	H-VNHQGDALSQLTSQL-OH
39	H-DALSQLTSQLQKNFQ-OH
40	H-LTSQLQKNFQAISSS-OH
41	H-QKNFQAISSSIAEIY-OH
42	H-AISSSIAEIYNRLEK-OH
43	H-IAEIYNRLEKVEADA-OH
44	H-NRLEKVEADAQVDRL-OH
45	H-VEADAQVDRLITGR-OH
46	H-KVEADAQVDRLITGR-OH

Appendix D: Proteomics

A folder containing data from the proteomics experiments can be accessed at the following address:

https://uob-my.sharepoint.com/:f:/g/personal/ss15846_bristol_ac_uk/EiMLoDzWuaRAtq71ZazCw3wBPFiVffcZP4ea9gbFE232Dg?e=VEnbAx